frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 13 November 2025
DOI 10.3389/feduc.2025.1676938

OPEN ACCESS

EDITED BY
Marko Hölbl,
University of Maribor, Slovenia

REVIEWED BY
Fady Alnajjar,
United Arab Emirates University,
United Arab Emirates
Hassan Abood,
Republic of Iraq Ministry of Oil, Iraq

*CORRESPONDENCE
Syaamantak Das
✉ syaamantak.das@iitb.ac.in

RECEIVED 31 July 2025
ACCEPTED 23 October 2025
PUBLISHED 13 November 2025

CITATION
Nagvekar PV, Das S and Iyer S (2025) Teaching log data analysis in Indian cybersecurity classrooms: a mixed-methods study of pedagogical challenges and learner difficulties. *Front. Educ.* 10:1676938. doi: 10.3389/feduc.2025.1676938

Teaching log data analysis in Indian cybersecurity classrooms: a mixed-methods study of pedagogical challenges and learner difficulties

Priya V. Nagvekar, Syaamantak Das* and Sridhar Iyer

Centre for Educational Technology, Indian Institute of Technology Bombay, Mumbai, India

Introduction: Log data analysis is a core competency in cybersecurity education, essential for investigating cyberattacks and identifying their root causes. However, teaching and learning this skill present distinct challenges in resource-constrained contexts such as India. Existing pedagogical approaches often fail to address the dual challenge of limited infrastructure and inadequate student preparation, resulting in persistent gaps between instructional intent and learner outcomes. This study aims to examine these gaps from both faculty and learner perspectives to understand systemic and cognitive barriers in cybersecurity education.

Methods: A mixed-methods design was adopted to explore pedagogical and cognitive challenges in teaching root cause analysis (RCA) through log data interpretation. First, a survey was administered to cybersecurity faculty members from diverse Indian institutions to identify systemic barriers, including insufficient prerequisite knowledge among students, insufficient infrastructure, and rigid curricula. Complementing this, an empirical study was conducted with undergraduate learners and industry experts. Participants performed RCA on simulated cyberattacks using logfiles and techniques, including the 5 Whys, fault trees, and attack trees. Comparative analysis focused on identifying reasoning patterns and problem-solving strategies across expert and novice groups.

Results: Survey data revealed consistent concerns among faculty regarding students' inadequate foundational knowledge, infrastructural limitations, and institutional rigidity that constrain pedagogical innovation. In the empirical phase, novice learners exhibited difficulties in technical interpretation, tendencies toward premature analysis termination, and several cognitive biases. In contrast, experts demonstrated structured reasoning, cross-functional integration, and methodical application of RCA techniques, highlighting a pronounced expert-novice divide.

Discussion: The findings indicate a significant misalignment between instructional objectives and students' preparedness, compounded by systemic institutional constraints. These insights underscore the need for curriculum redesign, targeted teaching strategies, and faculty development initiatives to better scaffold students' analytical reasoning in cybersecurity education. The study contributes to improving pedagogical practices in computing education within underrepresented and resource-limited contexts, offering a pathway to bridge the expert-novice divide in log data analysis training.

KEYWORDS

teaching log data, log data analysis, cybersecurity, student challenges, faculty practices

1 Introduction

Cybersecurity education has emerged as a strategic priority in the digital age, particularly as cyberattacks increasingly threaten individuals, institutions, and national infrastructure. Among the critical skills needed in cybersecurity is log data analysis, a core component of root cause analysis (RCA) that enables investigators to trace the origin, intent, and pathways of an attack. Despite its centrality to incident response and forensic investigations, log data analysis remains a complex skill to teach and learn, especially in undergraduate computer science (CS) programs in developing contexts like India.

In India, the teaching of cybersecurity is expanding, with new university-level programs and increased awareness of its relevance (Kant, 2023). Yet, the gap between academic training and industry requirements remains significant, particularly in the area of practical skills development, including log data interpretation and RCA. Faculty members often encounter obstacles such as insufficient student preparedness, infrastructure limitations, and rigid curricula, all of which hinder effective pedagogy. Simultaneously, students themselves face cognitive and conceptual difficulties in analyzing system logs, identifying vulnerabilities, and constructing RCA models like fault trees or attack trees. While both problems are individually acknowledged in literature, few studies have addressed them together from a systemic educational perspective within the Indian context.

India's cybersecurity education landscape reflects broader systemic issues common to developing higher education systems such as curricular rigidity, resource scarcity, and uneven faculty expertise. These contextual factors mirror those in other emerging economies, thereby positioning the present study as a transferable model for understanding pedagogical challenges in resource-constrained computing education.

This study addresses that gap by combining two distinct but complementary strands of inquiry. First, it presents results from a survey of 47 faculty members across Indian institutions who teach cybersecurity, highlighting institutional, curricular, and instructional challenges in teaching log data analysis. Second, it draws from an empirical study involving 24 novice learners (senior undergraduate CS students) and 3 industry experts who engaged in simulated RCA tasks using real-like log data from cyberattacks. By triangulating faculty perspectives with learners' process-level difficulties, we offer a comprehensive understanding of both instructional barriers and learner struggles in cybersecurity education in India.

We ask the following research questions:

- RQ1: What challenges and difficulties are faced by cybersecurity faculty when teaching log data analysis in India?
- RQ2: What difficulties do novice learners face while performing RCA using log data?

By answering these questions, this study contributes to the literature on cybersecurity education by identifying systemic misalignments between pedagogical intent and learner readiness. It also offers empirically grounded recommendations for curriculum reform, instructional design, and faculty development tailored to resource-constrained settings. Our findings are particularly relevant for computing education researchers, curriculum developers, and policymakers seeking to strengthen cybersecurity training in similar contexts.

2 Background and related work

Log data analysis is a foundational technique in cybersecurity, enabling professionals to trace malicious activities, determine the origin and progression of cyberattacks, and identify system vulnerabilities. Logs generated by firewalls, intrusion detection systems, servers, applications, and endpoints provide valuable digital traces that support RCA during incident response. Tools such as SIEM platforms[1] and structured approaches like the 5 Whys, fault tree analysis, fishbone diagrams, and attack trees help in systematically identifying how attacks unfold and what failures enabled them (Diogenes and Ozkaya, 2018,Johansen, 2020,Landauer et al., 2020, Jia et al., 2017).

In the context of computing education, the teaching of log data analysis is increasingly recognized as a necessary component of practical cybersecurity training. Prior studies (Liu et al., 2021, Hunt and Hill, 2015) have shown how analyzing log files using platforms like the ELK stack[2] can help learners visualize attack patterns and anomalies. Practical-based, case-study, and gamified approaches such as Capture the Flag (CTF) are increasingly adopted to support student engagement and learning outcomes in cybersecurity education (Švábenskỳ et al., 2021,McDaniel et al., 2016). These methods encourage active participation and support the application of theoretical knowledge to real-world problems.

However, the current literature on cybersecurity education is heavily skewed toward Western contexts. In developing countries like India, the pedagogical realities of teaching advanced cybersecurity concepts such as RCA and log data analysis are significantly different due to constraints in curriculum flexibility, lab infrastructure, and faculty expertise (Catota et al., 2019, Juyal et al., 2023, Mittal, 2024, Oladimeji et al., 2024). Indian cybersecurity instructors have reported (Mittal, 2024) that students often lack adequate prerequisite knowledge in key areas like networking, operating systems, and programming, skills critical for interpreting log data meaningfully.

Alongside institutional constraints, there are also learner-centered challenges. Research comparing experts and novices in RCA tasks has found that novices tend to exhibit premature closure, limited hypothesis generation, and single-cause attribution biases (Silva et al., 2015,Scheponik et al., 2016). They often struggle with integrating cross-domain knowledge and recognizing attack sequences from raw log entries. In contrast, experts are more systematic and holistic in their analysis, frequently cross-referencing log data with architectural understanding (Tovarňák et al., 2020, Silva et al., 2015). These differences have profound implications for how RCA and log interpretation should be taught.

Despite the growing recognition of these challenges, few studies have systematically analyzed both faculty teaching practices and

1 https://www.ibm.com/think/topics/siem

2 https://www.elastic.co/elastic-stack

learner difficulties together. Prior work has focused either on improving teaching environments (e.g., security labs (Topham et al., 2016); SEED Labs (Du, 2011) or on learner outcomes in competitive or simulated settings (McDaniel et al., 2016, Scheponik et al., 2016), but rarely triangulates both perspectives within a specific national context. There is also limited work exploring faculty development, although initiatives like "Security Across the Curriculum" have begun addressing this at a systemic level (Taylor and Kaza, 2016).

This study aims to fill this gap by combining two complementary lenses: a faculty survey across Indian institutions on teaching practices and constraints in log data analysis education, and a learner-focused empirical study comparing novice and expert performance on RCA tasks. Together, these insights offer a broader view of the structural and cognitive challenges that must be addressed to improve cybersecurity education in resource-constrained contexts like India.

3 Materials and methods

This study employed a sequential mixed-methods design consisting of two components: (Study 1) a national survey of cybersecurity faculty across Indian higher education institutions, and (Study 2) a performance-based controlled study comparing novice learners and expert practitioners on simulated log data analysis tasks. The faculty survey instrument included both Likert-scale and open-ended items, and was disseminated via professional mailing lists and institutional contacts. Responses (N = 47 complete) were analyzed using descriptive statistics and inductive thematic coding following the Braun and Clarke (2006) approach.

The novice–expert study involved 24 undergraduate students and 3 expert practitioners. Participants were given realistic log artifacts simulating common cybersecurity attack scenarios (DoS, XSS, Blind SQLi) and asked to complete RCA diagrams and verbal walkthroughs. Data sources included completed artifacts, time-on-task logs, and stimulated recall interviews. Qualitative analysis was conducted on learner outputs to identify reasoning patterns and errors, while comparisons between novice and expert performances were synthesized thematically.

All participants provided informed consent, and the study received ethics clearance from the Institutional Review Board of the host institute (IIT Bombay), approval number: IITB-IRB/2021/052.

4 Study design

4.1 Study 1: faculty survey on teaching log data analysis

This study aimed to examine the pedagogical challenges faced by Indian computer science faculty when teaching log data analysis as part of cybersecurity curricula. The survey focused on identifying institutional, curricular, and learner-related constraints, as well as instructional practices and improvement priorities from the faculty perspective.

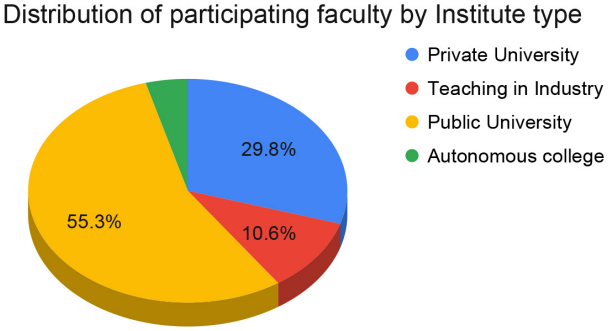

FIGURE 1
Distribution of participating faculty by Institute type.

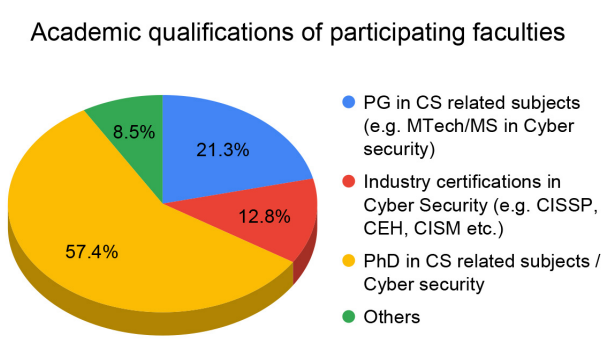

FIGURE 2
Distribution of academic qualifications of participating faculties.

4.1.1 Participant profile

The survey targeted faculty members who teach or have taught cybersecurity or network security courses in Indian colleges and universities. A curated national mailing list of 3,689 Computer Science faculty across 488 institutions was used for recruitment. This list represents a broad cross-section of Indian higher education institutions, including public and private universities, technical institutes, and affiliated colleges. Figure 1 shows the distribution of participants—Institute type wise. Figure 2 shows the distribution of academic qualifications of participating faculties. Figure 3 shows the distribution of Cyber Security Teaching Experience among participating faculties. Complete demographic summary for faculty participants is shown in Appendix Table 2.

The survey invitation,[3] sent via email on May 9, 2024, included a description of the study's objectives, voluntary participation terms, and assurances of confidentiality. The survey remained open for three months with periodic reminders.

A total of 47 complete responses were received. While this response rate is modest relative to the size of the mailing list, it is not atypical for open voluntary surveys in higher education, particularly on specialized topics such as cybersecurity pedagogy. Moreover, the respondent pool reflects a diverse and information-rich sample, capturing variation in institutional affiliation, teaching

3 Link for the survey - https://forms.gle/wy3XTPfzUPBwohr7A.

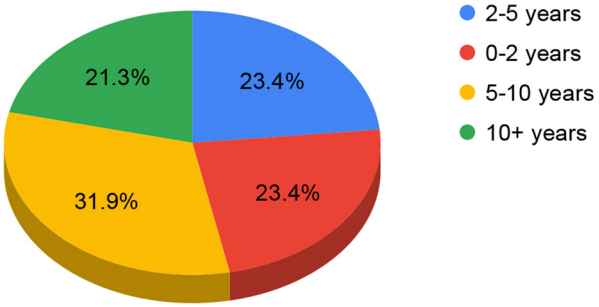

FIGURE 3
Distribution of participating faculty—Cyber Security Teaching Experience.

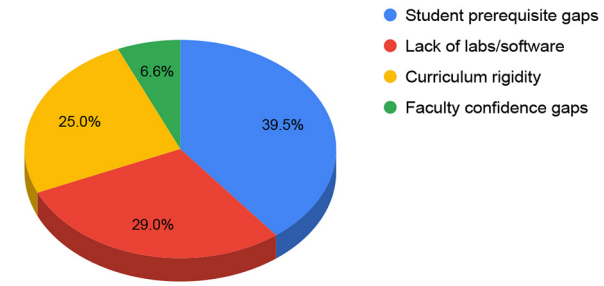

FIGURE 4
Faculty-reported challenges in teaching log data analysis.

experience, and academic qualification. These characteristics make the dataset suitable for exploratory mixed-methods analysis, consistent with prior education research in similar domains.

4.1.2 Survey instrument

The survey was administered in English using Google Forms. It was designed to capture a wide range of information across four domains:

1. Demographics: Academic qualifications, years of teaching experience, institution type (public/private), and teaching level (UG/PG).
2. Perceptions: Importance of log data analysis in cybersecurity education and its placement within the curriculum.
3. Teaching Practices: Methods used to teach log data analysis and perceived student challenges.
4. Improvement Priorities: Faculty preferences for addressing pedagogical barriers, including upskilling, curriculum reform, and infrastructure enhancement.

The instrument contained a mix of multiple-choice items, Likert-scale questions (1–5), and open-ended responses. Three faculty members from the authors' institution reviewed the survey for face validity, ensuring clarity and appropriate completion time.

4.1.3 Data analysis

A mixed-methods approach was used to analyze the survey data.

1. Quantitative analysis included descriptive statistics (mean, median, and standard deviation) for Likert-scale items and frequency analysis for categorical responses. Cross-tabulation was used to explore associations between variables such as faculty qualification and teaching practices.
2. Qualitative analysis was performed on open-ended responses using inductive coding (Braun and Clarke, 2006) to identify recurring themes related to teaching constraints, student preparedness, and faculty priorities.

4.2 Findings of study 1

4.2.1 Faculty consensus on the importance of log data analysis

Faculty respondents expressed strong consensus on the importance of teaching log data analysis for RCA in cybersecurity. Across all age groups, average agreement on its curricular importance was 4.29 out of 5. There was also support for introducing log data analysis in foundational cybersecurity courses (mean = 4.06; SD = 1.01), highlighting the perceived value of early exposure.

4.2.2 Answer to RQ1: challenges in teaching log data analysis

Three major barriers emerged:

1. Inadequate student prerequisites: Inadequate student prerequisites (62.5%), particularly in networking, programming, and operating systems, were the most frequently cited challenge.
2. Lack of specialized labs or software tools: Lack of specialized labs or software tools (46.9%) hindered practical demonstrations and exercises.
3. Curriculum rigidity: Curriculum rigidity (40.6%) limited faculty autonomy to introduce log-based activities or modify assessments.

Some faculty (11.6%) also reported their own knowledge or experience gaps, indicating a need for professional development. A detailed breakdown of prerequisite knowledge gaps showed that networking (N = 33) was the most commonly cited subject, followed by programming (N = 22), operating systems (N = 20), and database management systems (N = 12). Figure 4 shows faculty reported challenges in teaching log data analysis.

4.2.3 Teaching methods used

Faculty reported using four major mixes of approaches to teach log data analysis, including:

1. Textbook-based instruction (N = 8).
2. Hands-on classroom demonstrations using simulation data (N = 3).

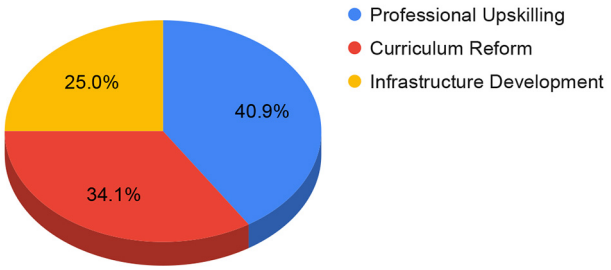

FIGURE 5
Faculty-preferred improvement priorities.

3. Real-life case studies from industry ($N = 3$).
4. Student-accessible datasets and lab exercises ($N = 13$).

However, curriculum constraints prevented some faculty ($N = 15$) from teaching log data analysis at all, despite acknowledging its importance.

Teaching methods varied by faculty qualification:

1. PhD holders and industry-certified faculty favored practical and interactive approaches ($N = 9$ combined).
2. Faculty from private institutions were more likely to adopt multiple teaching methods than their public-sector counterparts.

4.2.4 Faculty preferences for addressing challenges

The faculty identified diverse priorities for improvement, with professional upskilling, such as certifications like CompTIA Security+,[4] receiving the most first-preference votes ($N = 18$). Curriculum reform aimed at better integrating RCA and log analysis followed closely with 15 votes. Infrastructure development, including the establishment of dedicated labs, was recognized as a consistent secondary or tertiary priority. Preference patterns also varied according to faculty experience: less experienced faculty (0–2 years) emphasized the importance of curriculum reform, mid-career faculty (2–10 years) prioritized infrastructure improvements, while those with over ten years of experience focused on personal development and keeping up with evolving tools. Figure 5 shows faculty preferred improvement priorities in teaching log data analysis.

4.3 Study 2: novice and expert performance on RCA tasks

To complement the faculty perspective on teaching challenges, we conducted a learner-focused study to empirically investigate how novices and experts perform RCA using log data. This was a follow-up of an initial pilot study of similar nature (Nagvekar et al.). This study focused on identifying cognitive strategies, misconceptions, and structural patterns in reasoning that characterize novice performance, and contrasted them with expert practices. The findings reveal critical gaps in students' conceptual understanding and process-level reasoning that must be addressed through targeted pedagogy.

4.3.1 Participant details

The study included two participant groups:

1. Novice group: 24 senior undergraduate computer science students enrolled in cybersecurity or related elective courses. They had no formal training in RCA methods or professional cybersecurity experience.
2. Expert group: 3 industry professionals with 5–12 years of experience in cybersecurity roles such as Security Operations Center (SOC) analysts and incident responders. All had direct experience with log analysis and RCA in real-world attack scenarios.

4.3.2 Tasks and materials

Participants were given log data representing three types of simulated cyberattacks:

1. Simulated Log for Denial of Service (DoS) Attack
   This log file contains multiple, rapid HTTP GET requests from a single IP address (172.16.100.56) to the root endpoint (/) within a tight time window (1-second intervals). Each request receives a 499 response code (client closed request) and a content length of 0. The uniform user-agent and repeated requests indicate a flood-type DoS attack, aimed at overwhelming server resources or connections. This log was designed to help students identify temporal patterns and repeated request behavior typical of DoS scenarios. The sample log is given in Appendix 1.
2. Simulated Log for Cross-Site Scripting (XSS) Attack
   This log file traces typical web navigation followed by a reflected XSS payload submission. The user (192.168.61.155) interacts with multiple pages of a vulnerable application (bWAPP), eventually accessing xss_get.php with a URL containing encoded JavaScript payloads such as:

   ```
   <script>alert(1)</script> and
       alert(document.cookie)
   ```

   These appear in query parameters like firstname and lastname. The log showcases the injection vector and navigation trail that learners can use to identify input sanitization failures and client-side execution of malicious scripts. The sample log is given in Appendix 2.
3. Simulated Log for Blind SQL Injection
   This log simulates a time-based blind SQL injection against a vulnerable web form. The attacker from IP 192.168.61.155 submits GET requests with crafted SQL payloads like:

   ```
   username=admin or sleep(5)--
   ```

   and

   ```
   username=admin' or sleep(5)--
   ```

4 https://www.comptia.org/en/certifications/security/

TABLE 1 Summary of the log data-based tasks.

Log file	Attack type	Key features
Dos.txt	DoS	Repeated GET/requests, single IP, rapid sequence
XSS.txt	XSS	Script payloads in query params, form submission trail
SQL Injection.txt	Blind SQL Injection	Time-based payloads using sleep(5), or logic

The use of the sleep(5) function suggests a time-based strategy to infer database behavior without seeing direct output, a classic indicator of blind SQL injection. The repeated calls with slight variations help students identify payload evolution, parameter manipulation, and subtle attack behaviors. The sample log is given in Appendix 3.

They were asked to identify the root cause of each attack using one or more of the following RCA techniques:

1. 5 Whys (Tanimoto et al., 2023, Paterson, 2023).
2. Fault Tree Analysis (de Gusmão et al., 2018, Lallie et al., 2017, Kim et al., 2019).
3. Fishbone Diagram (Ishikawa) (Hellesen et al., 2018).
4. Attack Tree (Wang and Liu, 2014).

Novices received introductory tutorials on each method but were not formally trained. Experts could freely choose and apply any technique with which they were familiar. Table 1 shows a summary of the Log data-based tasks.

4.3.3 Data collection process

Each participant completed three log data analysis tasks (DoS, XSS, and Blind SQL Injection) in a randomized order within a 60-min session. Tasks were administered in a controlled lab environment with observation and think-aloud protocols. Analysis focused on structural coherence, causal depth, and terminology accuracy.

Data was collected through four methods:

- Participant-created RCA artifacts (e.g., diagrams, textual reasoning).
- Think-aloud notes and observations during task completion.
- Simulated recall interviews (Gass and Mackey, 2013) where participants reflected on their analysis decisions post-task.
- Time-on-task tracking to measure duration per activity.

All responses were anonymised, and consent was obtained before participation as per the authors' institutional IRB[5] process.

5 IITB-IRB/2021/052.

4.3.4 Data analysis

We used a three-phase multi-layered analysis framework combining:

1. Structural analysis: Completeness, coherence, and hierarchy of RCA outputs.
2. Content analysis: Correctness and specificity of nodes, technical terminology, and assumptions.
3. Thematic coding: Recurring errors, misconceptions, and reasoning strategies.

Expert responses served as a benchmark for comparison across the same three attack cases.

4.4 Findings of study 2

4.4.1 Answer to RQ2: novice difficulties

The novice group exhibited several recurring challenges:

1. Assumption bias: Many students inferred causes without grounding them in log evidence, often making leaps from symptoms to root causes.
2. Premature closure: Several participants stopped their analysis after identifying an initial issue, failing to pursue deeper systemic causes.
3. Lack of cross-functional reasoning: Students struggled to integrate concepts from operating systems, networking, and databases, even when such knowledge was crucial to interpreting log sequences.
4. Incorrect diagram structures: Fault trees and attack trees were often shallow, unbalanced, or contained contradictory logic flows.
5. Terminology gaps: Students used vague or incorrect terms (e.g., "*error happened*" instead of "*unauthorized SQL call*"), revealing a weak understanding of system components.

Time-on-task analysis showed high variability: novices spent longer on earlier tasks (mean ˜ 17 mins) but often rushed later ones, indicating fatigue or confusion.

4.4.2 Expert practices

In contrast, experts demonstrated four key characteristics:

1. Hierarchical and causal reasoning, constructing diagrams with consistent logic and well-labeled paths.
2. Attention to detail, identifying command sequences, IP traces, and HTTP request headers from logs.
3. Integrated knowledge across OS, database, and application layers.
4. Explanatory clarity in identifying not just what happened, but why it happened, and how to prevent it.

4.4.3 Implications from the expert–novice comparison

The analysis revealed a clear gap between how students are currently taught to reason through security problems and what expertise in RCA actually requires. Without scaffolding, novices

tended to either rely on surface-level log patterns or make uninformed guesses. Faculty-reported concerns about student preparedness (from Study 1) were strongly reflected in students' struggles during RCA tasks.

5 Discussion

The study offered a multi-stakeholder perspective on the challenges of teaching and learning log data analysis for cybersecurity education in India. By combining survey data from faculty members with an empirical study of student and expert performance on RCA tasks, we identify structural, instructional, and cognitive barriers that hinder effective skill development in this critical area. The findings demonstrate a pronounced misalignment between instructional goals, curricular implementation, and learner preparedness.

5.1 Faculty priorities and systemic constraints

The faculty survey revealed widespread agreement on the importance of teaching log data analysis early in cybersecurity curricula, especially for developing RCA capabilities. However, instructors cited several entrenched barriers: insufficient prerequisite knowledge among students, lack of dedicated lab infrastructure, and rigid curricula that restrict pedagogical autonomy. These findings suggest that even when faculty recognize the pedagogical importance of RCA and log data analysis, they often lack the institutional flexibility or resources to implement these practices meaningfully.

The variation in instructional practices ranges from textbook use to hands-on demos and case studies. This also highlights inconsistencies in teaching methods across institutions. Faculty from private universities reported using a broader mix of techniques than those in public institutions, but all groups faced curricular constraints that limited depth and authenticity in instruction. Notably, PhD-holding and industry-certified faculty were more likely to adopt interactive methods, underscoring the role of faculty expertise and exposure in shaping teaching approaches.

5.2 Novice learner struggles in RCA tasks

The learner study deepens these findings by showing how underprepared students struggle with both the structure and reasoning involved in RCA. Participants frequently displayed assumption bias, prematurely terminated their analysis, and produced diagrams with shallow or inconsistent logic. These difficulties map closely to the concerns raised by faculty: inadequate grounding in networking, programming, and systems-level thinking. Students were often unable to interpret log sequences in terms of system behavior or attacker intent, reflecting not only skill gaps but also conceptual and cognitive challenges in cybersecurity reasoning.

The contrast with expert performance was stark. Experts employed layered reasoning, incorporated multiple evidence streams, and demonstrated cross-domain integration. Their ability to articulate not just what happened, but why and how it could have been prevented, reveals the kind of sophistication that novices are currently not equipped to develop in typical classroom settings. This confirms the need for scaffolding instructional strategies that can gradually transition students from shallow to deep analysis.

These findings also contribute to the broader literature on expert–novice development. The struggles exhibited by novices in our study, such as premature closure, assumption bias, and fragmented reasoning, are consistent with prior models of novice performance in technical domains. Similar to findings in areas like clinical diagnostics and programming education, our results show that novices rely heavily on surface-level cues and lack the causal reasoning depth necessary for effective problem-solving. By contrast, experts demonstrated cross-functional integration and explanatory depth, reinforcing the importance of structured scaffolding in supporting students' progression from novice to more expert-like reasoning in cybersecurity contexts.

5.3 Misalignments between intent and practice

One of the most telling findings from this dual analysis is the contradiction between faculty aspirations and the realities of student capability and curricular design. Faculty generally support introducing log data analysis in foundational courses, but students lack the prerequisite knowledge to benefit from such instruction unless it is heavily scaffolded. Moreover, curricular inflexibility and the absence of labs or simulation environments further hinder meaningful engagement with log-based RCA tasks.

This misalignment echoes systemic issues in computing education in resource-constrained contexts. Even when pedagogical best practices are known, such as case-based learning, hands-on labs, or simulation-driven RCA. Faculty may not have the autonomy, resources, or training to implement them effectively. This gap contributes to the industry-reported observation that Indian cybersecurity graduates often lack practical readiness in key areas such as log interpretation and incident analysis.

5.4 Recommendations toward a holistic pedagogical approach

Our findings suggest that improving cybersecurity education in log data analysis requires interventions on multiple levels:

1. Curricular integration: RCA and log data analysis should be introduced in tandem with reinforcing foundational topics (e.g., networking, OS, databases), not as isolated modules.
2. Pedagogical scaffolding: Structured supports such as templates, partially filled diagrams, guided questions, and iterative feedback can help novices transition toward expert-like reasoning.
3. Faculty development: Instructors need access to continuing education in both technical content and pedagogy, particularly those without industry experience.

4. Infrastructure and tool access: Virtual labs, sandboxed environments, and annotated log datasets could offer scalable ways to expose students to authentic RCA experiences even where physical labs are lacking.

6 Implications for computing education

The findings of this study carry actionable implications for improving the teaching and learning of log data analysis and root cause reasoning in cybersecurity curricula, particularly in resource-constrained contexts like India. They offer guidance for educators, curriculum designers, and institutional leaders seeking to bridge the gap between theoretical knowledge and practical expertise.

6.1 Curriculum redesign for conceptual and procedural alignment

The widespread agreement among faculty on the importance of teaching log data analysis early in cybersecurity education suggests a need to restructure curricula to make this feasible. However, the recurring prerequisite gaps among learners in foundational areas such as networking, programming, and operating systems highlight the need for conceptual alignment. One implication is to adopt spiral curricula, where foundational concepts are revisited and deepened through increasingly complex RCA tasks.

Integrating RCA activities with system-level case studies across multiple semesters can help students gradually build the technical vocabulary and procedural fluency required for meaningful analysis. Institutions may also consider embedding RCA instruction into existing courses such as Operating Systems or Database Management, rather than isolating it within advanced electives.

6.2 Pedagogical scaffolding and cognitive support

Given the challenges students face in constructing logical and complete RCA diagrams, instructors should incorporate scaffolded pedagogical strategies. These may include:

- Partially completed diagrams that students extend or critique.
- Step-wise RCA exercises that guide learners through "what," "how," and "why" reasoning.
- Reflective prompts asking students to articulate alternative explanations.
- Progressive exposure to increasingly complex log data formats.

Scaffolding should be especially emphasized in foundational years, and gradually reduced as students gain proficiency. These supports can reduce cognitive overload while fostering structured analytical thinking.

6.3 Faculty capacity building and professional development

The study reveals that even faculty who value log data analysis often feel constrained by limited resources or curricular restrictions. Moreover, a subset of instructors reported gaps in their own confidence or training in teaching RCA. These insights point to the need for faculty development programs that focus not only on technical skills (e.g., interpreting SIEM logs or using attack tree tools), but also on pedagogical content knowledge on how to teach these skills effectively.

Workshops, micro-credentials, or modular courses on topics like "*Teaching Root Cause Analysis*" or "*Case-Based Cybersecurity Pedagogy*" could support instructors in both public and private institutions. Faculty mentorship models or cross-institutional teaching repositories may also help distribute high-quality teaching artifacts and reduce individual preparation burden.

These observations also foreground the role of pedagogical content knowledge (PCK) (Nilsson and Karlsson, 2019) in cybersecurity education. While faculty may possess technical expertise in RCA or log analysis, they often lack the resources or support to convert that expertise into pedagogically effective practices, particularly in contexts of curricular rigidity or infrastructural scarcity. Adopting a constructivist learning approach, where students iteratively build understanding through guided exploration of authentic log data and feedback-rich RCA tasks, can bridge this gap. Instructional designs that align with constructivist principles, such as problem-based learning, peer critique, and case reconstruction,can better support the cognitive development needed for meaningful learning in cybersecurity.

6.4 Infrastructure, tools, and simulation-based learning

Lack of specialized labs and software was a prominent barrier cited by faculty. Institutions with limited physical infrastructure can explore virtual labs and cloud-based simulation platforms that allow students to interact with log data in realistic environments. Open-source datasets, structured case banks, and sandboxed network emulators can approximate many of the benefits of on-premise labs at significantly lower cost.

Such tools could be further enhanced by annotated walkthroughs or embedded reflective questions that encourage learners to inspect logs with an RCA mindset. Collaboration with industry or government cybersecurity agencies to obtain anonymized log datasets could also help localize and contextualize learning.

6.5 Policy-level considerations for developing contexts

At a broader level, the study reinforces the need for education policy that balances theoretical depth with practical exposure in cybersecurity programs. Accrediting bodies and academic councils may consider including RCA and log data analysis as core learning

outcomes for cybersecurity tracks. Funding incentives for lab infrastructure or digital teaching materials, especially in public institutions, can help address institutional disparities.

Additionally, national faculty development missions could include cybersecurity pedagogy tracks, which would formally recognize and support educators engaging with the challenges documented in this study.

7 Limitations

This study examines faculty practices and student challenges in log data analysis within Indian cybersecurity education, highlighting several key insights despite some limitations. First, the faculty survey had a low response rate (47 out of 3,689), typical for specialized academic topics. However, the diverse sample enhances the validity of the qualitative analyses, even if it limits statistical generalizability. Second, the expert study featured a small group (N = 3), which limits comparisons between novice and expert performance. A larger expert sample in future studies would strengthen these benchmarks.Third, while using simulated log data was essential for ethical reasons, it may not fully capture the complexity of real-world security situations. Finally, while the study provides valuable insights into teaching practices, it lacks classroom observations or diverse evidence from teaching artifacts. Future research could build on these findings through longitudinal designs or ethnographic methods.

8 Conclusion

Log data analysis is a critical yet underdeveloped area in cybersecurity education, especially within resource-constrained contexts like India. This study combined faculty and learner perspectives to provide a holistic understanding of the challenges and opportunities in teaching RCA using log data in Indian computer science classrooms.

Through a national faculty survey, we identified widespread concerns around insufficient student prerequisites, lack of specialized infrastructure, and rigid curricula that limit pedagogical innovation. Simultaneously, our empirical study of novice learners revealed cognitive and conceptual difficulties in constructing RCA models from simulated log data, ranging from assumption bias and superficial reasoning to diagrammatic inconsistencies. The contrast with expert responses underscored the complexity of RCA as a skill that requires not only technical knowledge but also structured, cross-disciplinary thinking.

The triangulated findings highlight a systemic misalignment between curricular goals, institutional support, faculty capacity, and learner readiness. Addressing this gap requires multi-level interventions, including curriculum redesign to align RCA with prerequisite knowledge areas, scaffolded pedagogical strategies to support novice reasoning, faculty development programs focused on technical and instructional practices, and scalable simulation tools to support hands-on learning even in low-resource settings.

While the study is rooted in the Indian context, its implications extend to other developing nations seeking to modernize cybersecurity education and cultivate analytical competencies in their future security professionals. By foregrounding both structural and cognitive barriers, and by integrating faculty intentions with learner performance, this study contributes to the broader discourse on how to meaningfully teach complex cybersecurity practices in undergraduate computing education. In doing so, the study contributes to broader educational theory by extending models of PCK and constructivist learning into the domain of cybersecurity, an area still underrepresented in computing education research.

Future work should explore how integrated instructional interventions (e.g., case-based RCA modules, interactive simulations, and faculty co-design workshops) can improve student outcomes at scale. Expanding the dataset through longitudinal studies, diverse institutional contexts, and larger expert samples would further enrich the evidence base for reforming cybersecurity pedagogy globally.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Indian Institute of Technology Bombay. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

PN: Conceptualization, Data curation, Formal analysis, Methodology, Writing – review & editing. SD: Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing. SI: Conceptualization, Methodology, Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that Gen AI was used in the creation of this manuscript. The author(s) used Grammarly/Grammarly GO, an AI-enabled grammar and writing assistance tool, to support language clarity and proofreading during manuscript preparation. No generative AI tools were used for content generation, data analysis, or interpretation. All intellectual contributions, analyses, and interpretations presented in this manuscript are the author(s)' own.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2025.1676938/full#supplementary-material

References

Braun, V., and Clarke, V. (2006). Using thematic analysis in psychology. *Qual. Res. Psychol.* 3, 77–101. doi: 10.1191/1478088706qp063oa

Catota, F. E., Morgan, M. G., and Sicker, D. C. (2019). Cybersecurity education in a developing nation: the ecuadorian environment. *J. Cybersecur.* 5:tyz001. doi: 10.1093/cybsec/tyz001

de Gusmão, A. P. H., Silva, M. M., Poleto, T., e Silva, L. C., and Costa, A. P. C. S. (2018). Cybersecurity risk analysis model using fault tree analysis and fuzzy decision theory. *Int. J. Inf. Manage.* 43, 248–260. doi: 10.1016/j.ijinfomgt.2018.08.008

Diogenes, Y., and Ozkaya, E. (2018). *Cybersecurity-Attack and Defense Strategies*. Birmingham, UK: Packt Publishing.

Du, W. (2011). Seed: hands-on lab exercises for computer security education. *IEEE Secur. Priv.* 9, 70–73. doi: 10.1109/MSP.2011.139

Gass, S. M., and Mackey, A. (2013). *Stimulated Recall Methodology in Second Language Research*. London: Routledge. doi: 10.4324/9781410606006

Hellesen, N., Torres, H., and Wangen, G. (2018). Empirical case studies of the root-cause analysis method in information security. *Int. J. Adv. Secur.* 11, 60–79.

Hunt, R., and Hill, S. (2015). "Using security logs to identify and manage user behaviour to enhance information security," in *14th European Conference on Cyber Warfare and Security*, 111.

Jia, Z., Shen, C., Yi, X., Chen, Y., Yu, T., and Guan, X. (2017). "Big-data analysis of multi-source logs for anomaly detection on network-based system," in *2017 13th IEEE Conference on Automation Science and Engineering (CASE)* (IEEE), 1136–1141. doi: 10.1109/COASE.2017.8256257

Johansen, G. (2020). *Digital Forensics and Incident Response: Incident Response Techniques and Procedures to Respond to Modern Cyber Threats*. Birmingham, UK: Packt Publishing Ltd.

Juyal, T., Thapliyal, S., Garg, N., and Singh, D. (2023). "A study on cyber security and its challenges in india," in *International Conference on Information and Communication Technology for Intelligent Systems* (Springer), 151–159. doi: 10.1007/978-981-99-3761-5_15

Kant, R. (2023). Cyber-security awareness in india: how much students of higher education are aware? *Educ. Sci. Psychol.* 67, 59–72.

Kim, D., Kim, Y.-H., Shin, D., and Shin, D. (2019). Fast attack detection system using log analysis and attack tree generation. *Cluster Comput.* 22, 1827–1835. doi: 10.1007/s10586-018-2269-x

Lallie, H. S., Debattista, K., and Bal, J. (2017). An empirical evaluation of the effectiveness of attack graphs and fault trees in cyber-attack perception. *IEEE Trans. Inf. Foren. Secur.* 13, 1110–1122. doi: 10.1109/TIFS.2017.2771238

Landauer, M., Skopik, F., Wurzenberger, M., and Rauber, A. (2020). System log clustering approaches for cyber security applications: a survey. *Comput. Secur.* 92:101739. doi: 10.1016/j.cose.2020.101739

Liu, J.-C., Yang, C.-T., Chan, Y.-W., Kristiani, E., and Jiang, W.-J. (2021). Cyberattack detection model using deep learning in a network log system with data visualization. *J. Supercomput.* 77, 10984–11003. doi: 10.1007/s11227-021-03715-6

McDaniel, L., Talvi, E., and Hay, B. (2016). "Capture the flag as cyber security introduction," in *2016 49th Hawaii International Conference on System Sciences (HICSS)* (IEEE), 5479–5486. doi: 10.1109/HICSS.2016.677

Mittal, C. (2024). An empirical study on cybersecurity awareness, cybersecurity concern, and vulnerability to cyber-attacks. *Int. J. Sci. Res. Manag.* 12, 1144–1158. doi: 10.18535/ijsrm/v12i04.ec05

Nilsson, P., and Karlsson, G. (2019). Capturing student teachers' pedagogical content knowledge (PCK) using cores and digital technology. *Int. J. Sci. Educ.* 41, 419–447. doi: 10.1080/09500693.2018.1551642

Oladimeji, S., Egon, A., and Broklyn, P. (2024). Cybersecurity workforce development: Bridging the skills gap in the age of automation. *Available at SSRN 4904939*. doi: 10.2139/ssrn.4904939

Paterson, J. C. (2023). *Beyond the Five Whys: Root Cause Analysis and Systems Thinking*. New York: John Wiley Sons.

Scheponik, T., Sherman, A. T., DeLatte, D., Phatak, D., Oliva, L., Thompson, J., et al. (2016). "How students reason about cybersecurity concepts," in *2016 IEEE Frontiers in Education Conference (FIE)* (IEEE), 1–5. doi: 10.1109/FIE.2016.7757363

Silva, A., Emmanuel, G., McClain, J. T., Matzen, L., and Forsythe, C. (2015). "Measuring expert and novice performance within computer security incident response teams," in *International Conference on Augmented Cognition* (Springer), 144–152. doi: 10.1007/978-3-319-20816-9_15

Švábenský, V., Čeleda, P., Vykopal, J., and Brišáková, S. (2021). Cybersecurity knowledge and skills taught in capture the flag challenges. *Comput. Secur.* 102:102154. doi: 10.1016/j.cose.2020.102154

Tanimoto, S., Nakajima, R., Goromaru, H., Hatashima, T., and Kanai, A. (2023). "Risk countermeasures based on five whys analysis considering offensive security," in *2023 IEEE 12th Global Conference on Consumer Electronics (GCCE)* (IEEE), 640–642. doi: 10.1109/GCCE59613.2023.10315404

Taylor, B., and Kaza, S. (2016). Security injections@ towson: Integrating secure coding into introductory computer science courses. *ACM Trans. Comput. Educ.* 16, 1–20. doi: 10.1145/2897441

Topham, L., Kifayat, K., Younis, Y. A., Shi, Q., and Askwith, B. (2016). Cyber security teaching and learning laboratories: a survey. *Inf. Secur.* 35:51. doi: 10.11610/isij.3503

Tovarňák, D., Špaček, S., and Vykopal, J. (2020). Traffic and log data captured during a cyber defense exercise. *Data Brief* 31:105784. doi: 10.1016/j.dib.2020.105784

Wang, P., and Liu, J.-C. (2014). Threat analysis of cyber attacks with attack tree+. *J. Inf. Hiding Multim. Signal Proc.* 5, 778–788.