

ADDIS ABABA SCIENCE AND TECHNOLOGY UNIVERSITY

DEPARTMENT OF SOFTWARE ENGINEERING Master's Program (CEP)

Adaptive Online Examination Systems: Global Trends and the Ethiopian Context

Systematic Literature Review

Fidel Alemayehu Hailemariam

GSE0074/17

Submission Date: August, 2025
Submitted To: Dr. Girma

Domain Selection and Justification

An adaptive online examination system is a digital testing platform that changes the difficulty, sequence, or type of questions in real time based on a test taker's responses and performance. In addition to personalization, it integrates automated reporting mechanisms that provide detailed performance summaries, percentile calculations to compare a student's score with peers, and other grading methods such as scaled scoring, weighted topic analysis, and skill gap identification. These features enable faster feedback, fairer comparisons, and actionable insights for both learners and educators.

Currently, there are many institutions and organizations that need to evaluate students, candidates, or employees online. It is undeniable how online exams have skyrocketed in the last decade to meet academic and professional demands. The teaching methodology is evolving rapidly, resulting in virtual teaching thanks to the availability of advanced e-learning tools.

In Ethiopia, online examination systems have rapidly gained prominence in recent years, marking a significant shift in national assessment strategies. The adoption began in 2022 with the Graduate Admission Test and the National Graduate Admission Test, expanded in 2023 to include the National Exit Exam, and as of 2024, the National School Leaving (Matric) Examination is now conducted online. This trend reflects a growing recognition of the efficiency, scalability, and security offered by adaptive online examination systems, particularly in high-stakes national testing environments.

The developed adaptive exam system provides a wide variety of applications to students and lecturers due to its design and simple visual characteristics which do not require users to have any further guidance. By means of different methods, exams can be implemented, analyzed; and measurement and evaluation processes can be performed. In addition to the application of exams through an adaptive way, student monitoring and guidance as well as assessment of exam results, which is normally neglected due to its time-consuming structure but important in terms of achieving educational targets, can be more convenient.

Current Methods, Tools, and Data in Adaptive Online Examination Systems

Modern adaptive online examination systems rely on a combination of statistical models, machine learning algorithms, and secure delivery platforms to personalize assessments in real time.

One of the most widely used methods is Item Response Theory (IRT), including 1PL (**Parameter Logistic** model, and it describes how many parameters are used to model the relationship between a test-taker's ability and the probability of answering a question correctly.), 2PL, and 3PL models, which estimate a test taker's ability and select questions that match their skill level. This forms the basis for Computerized Adaptive Testing (CAT), where the difficulty and order of questions change dynamically as the test progresses. More advanced methods use

knowledge tracing models such as Deep Knowledge Tracing (DKT), Dynamic Key-Value Memory Networks (DKVMN), and Attentive Knowledge Tracing (AKT) to track how a learner's mastery evolves and predict future performance.

Machine learning plays a major role, with reinforcement learning and contextual bandit algorithms optimizing question sequencing policies to balance accuracy, difficulty, and content coverage. For maintaining integrity, AI-powered proctoring tools use computer vision techniques like face recognition (e.g., OpenCV, FaceNet), gaze tracking, and voice analysis to detect suspicious behaviors such as looking away, speaking during the exam, or switching screens.

Scoring mechanisms go beyond raw marks, incorporating percentile rank calculations to compare performance across cohorts, scaled scoring for standardized results, and weighted topic analysis to highlight strengths and weaknesses

From a development standpoint, frameworks such as TensorFlow, PyTorch, scikit-learn, R (catR, mirt), and Stan power the adaptive algorithms, while learning management systems like Moodle, Open edX, and Canvas integrate the exam delivery. Security is enforced through Safe Exam Browser (SEB) or Respondus LockDown Browser, and reporting dashboards built in Power BI, Tableau, or Metabase provide visual insights into performance trends, skill gaps, and percentile positions.

The data feeding these systems comes from multiple sources: calibrated question banks tagged with difficulty and topic metadata, interaction logs capturing response times and attempts, proctoring feeds from webcams and microphones for identity and cheating detection, and historical performance benchmarks used for norm-referenced grading. Together, these tools and datasets form the backbone of adaptive online examination systems.

Challenges

1. Infrastructure and Technical Limitations

Adaptive online exams rely on stable internet connections, reliable power, and device capabilities. Even in well-resourced contexts, high student loads, server outages, or cloud service failures can disrupt real-time item adaptation, proctoring, or reporting. Low-latency, synchronized delivery remains a technical challenge for large-scale deployments.

2. Algorithmic Fairness and Bias

AI-driven features such as facial recognition, gaze tracking, and NLP grading may exhibit biases across skin tones, genders, linguistic backgrounds, or cognitive abilities. Biases in adaptive item selection or scoring algorithms can affect measurement validity and fairness.

3. Privacy and Ethical Concerns

Continuous proctoring, interaction logging, and behavioral analytics raise global privacy

issues. Institutions must comply with data protection regulations and manage consent, retention policies, and ethical use of sensitive data.

4. **Item Bank Quality and Calibration**

Maintaining a well-calibrated item bank is challenging. Questions must be tagged with difficulty, discrimination, guessing parameters, and content areas. Without accurate calibration, adaptive algorithms may misestimate ability, producing unreliable scores.

5. **Measurement Validity and Transparency**

Shorter tests via CAT or knowledge tracing risk incomplete coverage of curricula or competencies. Furthermore, many adaptive systems lack explainable reasoning for question selection and scoring, reducing transparency for students and educators.

Opportunities

1. **Efficiency and Precision**

Adaptive systems reduce test length while maintaining precision, enabling faster assessment cycles with minimal fatigue. Real-time skill estimation allows personalized learning feedback and targeted remediation.

2. **Risk-Adaptive Proctoring**

AI can monitor for suspicious behavior (face recognition and movement, voice analysis) and escalate only when necessary, reducing surveillance fatigue and improving student trust.

3. **Edge Computing and Offline Capability**

On-device processing and offline-first designs reduce reliance on constant connectivity, enhancing reliability for global users in varying network conditions.

4. **Learning Analytics and Reporting**

Advanced dashboards and reporting mechanisms allow percentile ranking, mastery analysis, skill-gap identification, and predictive modeling, supporting both formative and summative assessment.

5. **Inclusive and Accessible Design**

Mobile-first interfaces, screen-reader compatibility, adjustable font sizes, and culturally neutral content enable broader participation, including students with disabilities.

Research Gaps in Adaptive Online Examination Systems

1. **Benchmarking under Stress Conditions**

Most current studies focus on ideal conditions with stable internet, sufficient server capacity, and reliable devices. However, in real-world global deployments, systems often face high-load scenarios, such as thousands of simultaneous test-takers, low-latency requirements, or partially offline conditions. There is a lack of research quantifying the

trade-offs between system performance, reliability, and assessment integrity under these stress conditions.

2. **Algorithmic Fairness and Bias Audits**

Adaptive exams increasingly rely on AI for proctoring, item selection, and automated scoring. Yet, there are few standardized frameworks to audit or measure bias across gender, ethnicity, disability, or language groups. For example, facial recognition may underperform for certain skin tones.

3. **Privacy-Preserving Adaptivity**

Adaptive systems collect sensitive data, including responses, behavior logs, webcam streams, and NLP-processed answers.

4. **Explainable Adaptive Systems**

Most adaptive online exams provide results and item selection without explaining why specific questions were presented or why certain scores were assigned. This lack of transparency reduces trust among students and educators.

5. **Open Simulators and Standardized Item Banks**

There is a scarcity of reproducible global datasets, simulators, and standardized item banks for CAT and knowledge-tracing research.

6. **Exam Security Gaps and Cheating Styles**

Despite proctoring technologies, online exams remain vulnerable to various cheating methods. Current research gaps include:

Screen manipulation: Switching, mirroring, or sharing screens to access unauthorized content.

AI-assisted cheating: Using large language models or generative AI to produce answers in real time.

Collusion: Coordinated answer sharing among multiple test-takers.

Proxy test-taking: Hiring someone else to take the exam using stolen credentials.

System exploitation: Timing attacks, exploiting adaptive item selection patterns, or manipulating response inputs to game scoring algorithms.

Research is limited on comprehensive detection and prevention frameworks that combine AI analytics, behavioral modeling, and risk-adaptive proctoring.

Ethiopian Context Analysis: Adaptive Online Examination Systems

1. Local Challenges

- **Internet & Power Issues:** Low bandwidth, unstable connections, and frequent outages disrupt exams and risk incomplete submissions.
- **Low Digital Literacy:** Limited familiarity with platforms increases errors and test anxiety.
- **Weak Infrastructure:** University systems struggle with large-scale, simultaneous testing, causing slowdowns and crashes.
- **Security & Cheating Risks:** Vulnerable to unauthorized access, screen sharing, collusion, and AI-assisted cheating.
- **Limited Adaptivity:** Most systems lack real-time adaptive item selection, advanced proctoring, and diverse question formats beyond MCQs.
- **Multilanguage support**
- **Not getting enough research papers and information in Ethiopia about online examination.**

2. Current Strengths

1. **Secure Exam Delivery**
Exams like the Graduate Admission Test (GAT) and National Graduate Admission Test (NAGT) use one-time login sessions and secure browsers, ensuring that test-takers cannot access the system multiple times or navigate away during the exam.
2. **Adaptive Reporting and Analytics**
These exams feature percentile-based reporting, allowing students and institutions to compare performance across the cohort. This provides clear insights into relative achievement and can inform targeted interventions.
3. **User-Friendly and Interactive Interfaces**
The platforms support interactive UIs, allowing students to flag questions for review and navigate easily between sections. The systems are designed to be easy to use, fast, and accessible, improving the test-taking experience.
4. **Foundations for Adaptivity**
While full adaptive item selection is limited, these systems demonstrate a baseline of digital exam management, including timed assessments, instant scoring for multiple-choice items, and secure proctoring workflows.

Methodology

This systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to ensure transparency, reproducibility, and rigor in the review process.

1. Search Strategy

A comprehensive literature search was conducted across major academic databases, including:

- IEEE Xplore
- ACM Digital Library
- Supplementary searches in Google Scholar, emerald, researchGate.

Keywords and Search Strings:

Search queries were formulated using Boolean operators to combine terms related to adaptive learning and proctoring, for example:

("adaptive web personalization" OR "computerized adaptive testing" OR "adaptive assessment" OR "online Testing" OR "online Exam ") OR Online Exam Questions)

AND ("online examination" OR "remote proctoring" OR "exam monitoring")

Searches were customized for each database to ensure maximum retrieval.

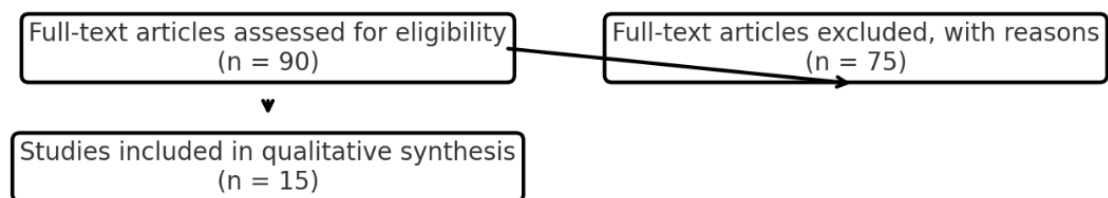
2. Inclusion and Exclusion Criteria

Inclusion:

- Peer-reviewed journal articles, conference papers, and systematic reviews.
- Studies published between 2020 and 2025.
- Publications focusing on adaptive learning systems, adaptive web personalization, or AI/ML-driven online examination systems.
- Research applicable to Ethiopian contexts included if available.

Exclusion:

- Non-academic blog posts, white papers, and non-peer-reviewed articles.
- Studies without English or Amharic translations.
- Publications unrelated to adaptive personalization in assessment or monitoring.



Overview

Adaptive online exams dynamically adjust question difficulty for greater accuracy and efficiency. While offering strengths like personalization, security features, and scalability, they face significant challenges including security vulnerabilities, infrastructure dependency, fairness concerns, and privacy risks.

Opportunities exist in AI-enhanced proctoring, hybrid offline-online models, and mobile-first design. Critical research gaps remain in adapting to low-resource settings, detecting AI-assisted cheating, ensuring cultural fairness, and developing transparent algorithms.

References

- [1] Q. Liu *et al.*, “Survey of Computerized Adaptive Testing: A Machine Learning Perspective,” *arXiv preprint*, 2024. [\[2404.00712\] Survey of Computerized Adaptive Testing: A Machine Learning Perspective](#)
- [2] A. Abdrashitov *et al.*, “A Systematic-Narrative Review of Online Proctoring Systems,” *Open Praxis*, 2025. [A Systematic-Narrative Review of Online Proctoring Systems and a Case for Open Standards | Open Praxis](#)
- [3] G. Whitelock *et al.*, “Digital proctoring in higher education: a systematic literature review,” *Int. J. Educational Management*, 2023. [Digital proctoring in higher education: a systematic literature review | International Journal of Educational Management | Emerald Publishing](#)
- [4] “Fast, smart, and adaptive: using machine learning to optimize assessments with CAT,” *Scientific Reports*, 2025. <https://www.nature.com/articles/s41598-025-91086-w>
- [5] NeurIPS/ACM paper/poster, “Computerized Adaptive Testing via Collaborative Ranking,” *NeurIPS/ACM Web Conf. proceedings*, 2024. <https://neurips.cc/virtual/2024/poster/96601>
- [6] M. Turani *et al.*, “Deep Learning for Online Exam Cheating Detection,” *Education Sciences (MDPI)*, 2023. <https://www.mdpi.com/2078-2489/14/3/195>
- [7] Safe Exam Browser project — technical documentation and institutional case studies, SEB (various institutional reports). <https://www.emerald.com/ijem/article/38/1/265/1228105/Digital-proctoring-in-higher-education-a>
- [8] X. Yang *et al.*, “iExam: Online Exam Monitoring and Analysis based on Face Detection,” *conference/arXiv*, 2022.
- [9] Iqbal *et al.*, “Enhancing Online Exam Security: Deep Learning Algorithms for Cheating Detection,” *conference/journal*, 2023. [Web-Based Exam Management System: Improving Efficiency and | Course Hero](#)
- [9] C. Hunduma, “Undergraduate Students' Perception of Exit Examination at Haramaya University”, 2023. [\(PDF\) Undergraduate Students' Perception of Exit Examination at Haramaya University](#)
- [10] AmaSQuAD: “A Benchmark for Amharic Extractive Question Answering,” *arXiv*, 2025 useful for low-resource QA and Amharic NLP. <https://www.mdpi.com/2078-2489/14/3/195>
- [11] Y. Berehanu, Y. Abera, Y. Abate, F. Alemayehu, “*Smart Exam Monitoring System in Ethiopia*,” 2023.
- [12] Iqbal *et al.*, “*Enhancing Online Exam Security: Deep Learning Algorithms for Cheating Detection*,” *conference/journal paper* (2023). [Enhancing Online Exam Security: Deep Learning Algorithms for Cheating Detection | Request PDF](#)
- [13] S. (Author), “Digital proctoring in higher education: a systematic literature review,” *International Journal of Educational Management*, 2023. [Emerald](#)
- [14] *Digital proctoring in higher education: a systematic literature review*, *Int. J. Educational Management*, 2023. [Emerald](#)
- [15] Educational Assessment and Examination Services, 2022 [EAES - Educational Assessment and Examination Services](#)