

Implementing Computer Adaptive Testing for High-Stakes Assessment: A Shift for Examinations Council of Lesotho

Enviado: 9 de septiembre de 2024 / Aceptado: 12 de septiembre de 2024 / Publicado: 9 de diciembre de 2024

MUSA ADEKUNLE AYANWALE

Department of Educational Foundations, National University of Lesotho, Lesotho
ma.ayanwale@nul.ls

 [0000-0001-7640-9898](https://orcid.org/0000-0001-7640-9898)

JULIA CHERE-MASOPHA

Department of Educational Foundations, National University of Lesotho, Lesotho
juliachere@gmail.com

 [0000-0002-8279-2640](https://orcid.org/0000-0002-8279-2640)

MAPULANE MOCHEKELE

Department of Educational Foundations, National University of Lesotho, Lesotho
mapulanemochekele@gmail.com

 [0009-0004-6384-6844](https://orcid.org/0009-0004-6384-6844)

MALEBOHANG CATHERINE MORENA

Department of Educational Foundations, National University of Lesotho, Lesotho
cmorena2008@yahoo.com

 [0000-0002-5347-474X](https://orcid.org/0000-0002-5347-474X)

DOI 10.24310/ijne.14.2024.20487

ABSTRACT

We examine the feasibility of implementing Computer Adaptive Testing (CAT) for high-stakes assessments in Lesotho, specifically through the Examination Council of Lesotho (ECoL). CAT, a cutting-edge testing method, enhances precision and efficiency by adjusting test items in real-time based on an examinee's ability. While CAT has gained widespread global adoption, its implementation in developing countries presents significant challenges, particularly regarding infrastructure, expertise, and resource limitations. Our research addresses a critical gap in understanding how CAT can be effectively integrated

RESUMEN

Implantación de las pruebas adaptativas por ordenador para evaluaciones de alto nivel: Un cambio para el Consejo de Exámenes de Lesotho

Examinamos la viabilidad de la aplicación de las Pruebas Adaptativas por Ordenador (CAT) a las evaluaciones de alto nivel en Lesoto, concretamente a través del Consejo de Exámenes de Lesoto (ECoL). El CAT, un método de evaluación de vanguardia, mejora la precisión y la eficacia ajustando los elementos de la prueba en tiempo real en función de la capacidad del examinado. Aunque el CAT se ha adoptado de forma generalizada



into educational systems in developing contexts such as Lesotho. We conducted a literature review across five databases—Google Scholar, ERIC, PsycINFO, JSTOR, and PubMed—retrieving 48 studies published between 2013 and 2023. Of these, 18 studies met our inclusion criteria, focusing on CAT’s advantages, challenges, and real-world applications in educational assessments. We applied thematic analysis to identify key benefits and barriers, focusing on precision, fairness, and infrastructure requirements. We outline the essential stages of CAT development, including feasibility studies, item bank creation, pretesting and calibration, specification determination, and live CAT publication. We highlight critical tasks such as using Monte Carlo simulations to validate CAT feasibility and developing a robust item bank calibrated with Item Response Theory (IRT). We also address challenges like building a robust technological infrastructure, providing comprehensive stakeholder training, and securing adequate funding. We emphasize the importance of continuous evaluation and stakeholder engagement to ensure CAT’s successful implementation and sustainability. Global trends indicate growing adoption, driven by advancements in psychometrics and technology. CAT has the potential to offer more equitable and accurate assessments, making it a promising solution to improve educational outcomes in Lesotho.

Key words: Assessment technology, computer adaptive testing (CAT), educational measurement, examination council of Lesotho, high-stakes assessments, item response theory (IRT).

en todo el mundo, su aplicación en los países en vías de desarrollo presenta importantes retos, especialmente en lo que respecta a la infraestructura, la experiencia y las limitaciones de recursos. Nuestra investigación aborda una laguna crítica en la comprensión de cómo la TAO puede integrarse eficazmente en los sistemas educativos de contextos en desarrollo como el de Lesoto. Llevamos a cabo una revisión de la literatura a través de cinco bases de datos—Google Scholar, ERIC, PsycINFO, JSTOR y PubMed—recuperando 48 estudios publicados entre 2013 y 2023. De estos, 18 estudios cumplieron con nuestros criterios de inclusión, centrándose en las ventajas, desafíos y aplicaciones del mundo real de CAT en evaluaciones educativas. Se aplicó un análisis temático para identificar los principales beneficios y obstáculos, centrándose en la precisión, la equidad y los requisitos de infraestructura. Esbozamos las etapas esenciales del desarrollo de un CAT, incluidos los estudios de viabilidad, la creación de un banco de ítems, las pruebas preliminares y la calibración, la determinación de las especificaciones y la publicación del CAT en vivo. Destacamos tareas críticas como el uso de simulaciones Monte Carlo para validar la viabilidad de los CAT y el desarrollo de un banco de ítems robusto calibrado con la Teoría de Respuesta al Ítem (TRI). También abordamos retos como la creación de una infraestructura tecnológica sólida, la formación exhaustiva de las partes interesadas y la obtención de financiación adecuada. Hacemos hincapié en la importancia de la evaluación continua y la participación de las partes interesadas para garantizar el éxito de la aplicación y la sostenibilidad del CAT. Las tendencias mundiales indican una creciente adopción, impulsada por los avances en psicometría y tecnología. El CAT tiene el potencial de ofrecer evaluaciones más equitativas y precisas, lo que lo convierte en una solución prometedora para mejorar los resultados educativos en Lesotho.

Palabras clave: Tecnología de la evaluación, pruebas adaptativas por ordenador (CAT), medición educativa, consejo examinador de Lesoto, evaluaciones de alto nivel, teoría de la respuesta al ítem (IRT).

1. INTRODUCTION

Computer Adaptive Testing (CAT) represents a significant advancement in the field of educational assessment, offering a more dynamic and individualised approach to testing compared to traditional fixed-form or paper-pencil assessments. The core principle of CAT is its ability to adjust the difficulty of test items in real-time based on the examinee's performance, thereby providing a more precise measurement of their abilities (Ogunjimi et al., 2021; Thompson, 2023). This adaptability is primarily achieved through the use of Item Response Theory (IRT), which allows for the creation and calibration of items that can accurately gauge an examinee's ability across a wide range of proficiency levels (Han, 2018b). The implementation of CAT in educational settings has been facilitated by advances in technology and psychometrics. Modern CAT systems utilise sophisticated algorithms to select the most appropriate items for each examinee, ensuring that the test remains both challenging and fair regardless of the examinee's ability level (Thompson & Weiss, 2009, 2011). This process not only enhances the precision of ability estimates but also reduces the number of items required to achieve a reliable assessment, thereby decreasing the overall test duration (Thompson, 2023; Thompson, 2009). The efficiency of CAT makes it particularly suitable for high-stakes assessments (such as the one administered by Examination Council of Lesotho, ECoL) where accuracy and fairness are critical (Viswanandhne & Nandakumar, 2017).

Furthermore, the development of CAT has been driven by the need to address the limitations of traditional assessment methods. Fixed-form tests often fail to account for the diverse abilities of examinees, leading to issues such as test fatigue for higher-ability students and disengagement for lower-ability ones (Meijer & Nering, 1999). By continuously adapting to the examinee's performance, CAT provides a more engaging and tailored testing experience, which can lead to more accurate assessments and better outcomes for students (Ogunjimi et al., 2021). The ECoL is the authority responsible for developing, overseeing, and maintaining a robust assessment system for Basic Education that meets Lesotho's educational requirements. ECoL administers various assessments, including the Grade 7 End-of-Level Assessment and the Grade 11 public examinations for the Lesotho General Certificate of Secondary Education (LGCSE), a qualification endorsed by the National University of Lesotho (NUL) for quality assurance. Additionally, ECoL conducts the biennial Lesotho National Assessment of Educational Progress (LNAEP) Survey in Grades 4 and 6, monitoring the progress of the country's education system on behalf of the Ministry of Education and Training (MoET) (Ayanwale et al., 2022; examsCouncil.org.ls). Traditional assessment methods employed by ECoL have limitations in accurately assessing examinees' abilities and adapting to their differences. This



study aims to evaluate the feasibility of introducing CAT in Lesotho, assessing its potential benefits, challenges, and overall impact on the educational system.

2. METHOD

We systematically reviewed the literature on the implementation and challenges of Computer Adaptive Testing (CAT), focusing on its applicability in high-stakes assessments in educational contexts, particularly within the Examination Council of Lesotho (ECoL). Our methodology involved multiple stages, including keyword selection, database search, and applying inclusion and exclusion criteria to identify relevant studies. We searched our literature using five electronic databases: Google Scholar, ERIC (Education Resources Information Center), PsycINFO, JSTOR, and PubMed. These databases were selected to provide a broad and comprehensive search across education, psychology, and assessment-related research fields. We retrieved 20 articles from Google Scholar, 13 from ERIC, 5 from PsycINFO, 7 from JSTOR, and 3 from PubMed, yielding a total of 48 relevant studies. To locate articles discussing the advantages, implementation, and challenges of CAT, we used a combination of specific keywords, including “Computer Adaptive Testing (CAT),” “Advantages of CAT in high-stakes assessments,” “Challenges of CAT implementation,” “Item Response Theory (IRT) and adaptive testing,” “Feasibility of CAT in developing countries,” “Educational technology and adaptive testing,” “CAT and fairness in assessments,” “Psychometrics and computerized testing,” and “CAT infrastructure challenges in education.” Boolean operators (AND, OR) were employed to refine the search results and ensure relevance to the study’s focus.

We limited our search to peer-reviewed articles published between 2013 and 2023 to capture the most recent advancements in CAT technology and its implementation in educational settings. Our review prioritized research articles and books that provided empirical data on CAT and its use in high-stakes assessments. For inclusion, we selected studies focusing on CAT, its advantages, and implementation challenges in high-stakes assessment environments. Articles that addressed CAT in developing countries or provided empirical evidence of its feasibility were also prioritized. We excluded studies that did not focus on CAT or dealt solely with low-stakes assessments, as well as those without peer-reviewed credibility or lacking empirical data. After identifying relevant articles, we screened them by title and abstract to ensure alignment with our study objectives. We then conducted a full-text screening of the selected articles to extract detailed information on CAT’s benefits, challenges, and real-world applications. This process resulted in the final selection of 18 articles for analysis from

different databases such as: Google Scholar (7 articles), ERIC (5 articles), PsycINFO (2 articles), JSTOR (2 articles), and PubMed (2 articles). We extracted and analyzed data from these articles, focusing on key themes such as CAT's precision, efficiency, fairness, and infrastructure requirements. Special attention was given to the potential of CAT in enhancing assessment practices in developing countries, and our analysis employed critical discourse techniques to offer a balanced discussion of both the advantages and challenges of CAT in high-stakes assessments.

3. BENEFITS OF COMPUTER ADAPTIVE TESTING

One of the primary benefits of CAT is its ability to enhance the precision and accuracy of assessments. By selecting items that are most appropriate for the examinee's current ability level, CAT ensures that each question provides maximum information about the test-taker's abilities. This leads to more reliable scores with fewer items compared to traditional testing methods (Thompson, 2009; Thompson & Weiss, 2011). The adaptive nature of CAT allows it to adjust in real-time, providing a personalized testing experience that can better gauge an individual's knowledge and skills (Han, 2018; Han, 2012; Tsaousis et al., 2021). This precision is particularly valuable in high-stakes testing environments, where the accuracy of test results can significantly impact educational and career opportunities. CAT also significantly reduces test length, which decreases test fatigue and anxiety among examinees. Studies have shown that adaptive tests can reduce the number of questions needed to achieve the same level of measurement precision by 30% to 50% (Han, 2018). This reduction in test length not only makes the testing experience more manageable for students but also allows for more efficient use of time and resources in educational settings. The ability to complete assessments more quickly without compromising accuracy is a key advantage of CAT over traditional fixed-form tests (Ogunjimi et al., 2021; Oladele et al., 2020). Moreover, CAT offers improved test security and fairness. Since each examinee receives a unique set of questions tailored to their ability level, it becomes much more difficult to share answers or engage in other forms of cheating (Thompson, 2023). This individualized approach also helps to reduce the chances of item overexposure, which can be a significant concern in traditional testing methods where the same items are reused across multiple test administrations (Thompson & Weiss, 2009). The use of sophisticated algorithms to select and deliver test items ensures that each examinee's test is unique, thereby enhancing the overall security and integrity of the assessment process (Mujtaba & Mahapatra, 2020).



The fairness of CAT is another notable benefit. Traditional fixed-form tests often disadvantage certain groups of examinees by including items that are either too easy or too difficult relative to their abilities. In contrast, CAT adjusts the difficulty of items in real-time, ensuring that each examinee is presented with questions that are appropriately challenging (Leroux et al., 2013; Meijer & Nering, 1999). This adaptive approach not only provides a more accurate measure of an examinee's abilities but also creates a more equitable testing environment. The immediate feedback provided by CAT can further enhance learning outcomes by allowing educators to quickly identify and address areas where students may need additional support (Han, 2012).

4. CHALLENGES OF IMPLEMENTING CAT

Despite its numerous benefits, implementing CAT presents several significant challenges. One of the primary obstacles is the need for a robust technological infrastructure. Effective CAT deployment requires reliable computer hardware, secure internet connectivity, and sophisticated software capable of administering and scoring adaptive tests in real-time (Seo, 2017; Wainer et al., 2000). This can be particularly challenging in developing regions, such as Lesotho, where access to technology and digital literacy may be limited. Ensuring that all schools have the necessary equipment and connectivity is a critical first step in implementing CAT (Ayanwale & Ndlovu, 2022; Stepanek & Martinkova, 2020). Additionally, developing a comprehensive item bank is critical for the success of CAT. Items must be meticulously calibrated using Item Response Theory (IRT) to ensure they accurately measure the intended abilities across a range of difficulty levels (Meijer & Nering, 1999). The process of creating and validating such an item bank is time-consuming and resource-intensive. It requires expert input from psychometricians, subject matter experts, and software developers (Thompson, 2023). Moreover, maintaining the security and integrity of the item bank is essential to prevent item overexposure and ensure that the adaptive testing process remains fair and accurate (Thompson & Weiss, 2011).

Another significant challenge is ensuring digital literacy among both test-takers and administrators. In many developing countries, including Lesotho, there may be limited experience with computer-based testing. Comprehensive training programs are needed to familiarize educators, students, and administrators with the CAT system and its benefits (Han, 2018a; Han, 2016). This training must cover not only the technical aspects of using CAT but also the pedagogical implications of adaptive testing and how it can be integrated into the broader

educational framework. Furthermore, stakeholder buy-in is essential for the successful implementation of CAT. Educators, administrators, and students must be adequately trained and convinced of the benefits of CAT over traditional assessment methods. Addressing concerns about the fairness and reliability of adaptive testing, as well as ensuring that all stakeholders are comfortable with the technology, is vital for a smooth transition. This includes engaging with policymakers and education leaders to secure the necessary support and resources for the implementation of CAT. Financial constraints also pose a challenge. The initial costs of setting up the technological infrastructure, developing the item bank, and providing training can be substantial. While CAT can lead to cost savings in the long run through more efficient testing processes, the upfront investment required can be a significant barrier. Securing funding from government sources, international organizations, or private sector partnerships may be necessary to overcome this hurdle. By addressing these challenges through careful planning and resource allocation, ECoL can successfully implement CAT and reap its many benefits. This includes developing a phased implementation plan that allows for gradual integration and continuous evaluation to ensure the system meets the needs of all stakeholders.

5. GLOBAL TRENDS AND APPLICATIONS OF CAT

CAT has been widely adopted across various educational and professional settings globally, showcasing its versatility and effectiveness. In the United States, CAT is extensively used in high-stakes examinations for health professionals, such as the National Council Licensure Examination (NCLEX) for nurses and the Graduate Record Examinations (GRE) for graduate school admissions. These implementations have demonstrated significant reductions in test length and improvements in measurement precision, making CAT a preferred choice for many licensing and certification bodies (Han, 2018b). The success of CAT in the United States has influenced other countries to explore and implement adaptive testing in their educational systems. For instance, the United Kingdom has integrated CAT into certain university entrance examinations, while countries like Korea and Australia are exploring its potential for national licensing exams in various professional fields (Han, 2018b; Ogunjimi et al., 2021). These international examples underscore the adaptability of CAT to different educational contexts and the benefits it brings in terms of efficiency and accuracy (Ayanwale & Ndlovu, 2022; Meijer & Nering, 1999). The widespread adoption of CAT is facilitated by continuous advancements in psychometric research and technology, which enhance the stability and reliability of adaptive tests (Viswanandhne & Nandakumar, 2017).

In developing countries, the adoption of CAT is beginning to gain attraction as well. For example, several African nations such as Nigeria, South Africa, Egypt, to mention few are considering CAT for their national assessments to improve the quality and fairness of their educational evaluations (Ogunjimi et al., 2021). The potential for CAT to provide more accurate and equitable assessments is particularly appealing in these contexts, where educational disparities are often pronounced. However, the implementation of CAT in developing countries requires addressing unique challenges related to technological infrastructure, digital literacy, and resource allocation. Simulation tools like SimulCAT play a crucial role in evaluating and optimizing CAT implementations by allowing test developers to model and analyze different testing scenarios before live deployment (Han, 2012). These tools help ensure that CAT systems are tailored to meet the specific needs of different educational and professional contexts, thereby maximizing their effectiveness and efficiency. The use of simulation studies in the development of CAT systems has become a standard practice, providing valuable insights into item selection algorithms, test length optimization, and the management of item exposure (Oladele et al., 2022; Wainer et al., 2000). Moreover, the continuous improvement of CAT algorithms and methodologies contributes to its global success. Advances in item response theory (IRT) and machine learning have enhanced the precision and adaptability of CAT, making it a more robust and reliable assessment tool. These technological advancements enable CAT to provide more accurate measurements of examinees' abilities, even in diverse and heterogeneous populations (Mujtaba & Mahapatra, 2020; Rice et al., 2022). As a result, more countries and educational institutions are recognizing the potential of CAT to transform their assessment practices.

6. USE OF CAT IN DEVELOPING COUNTRIES: FOCUS ON LESOTHO

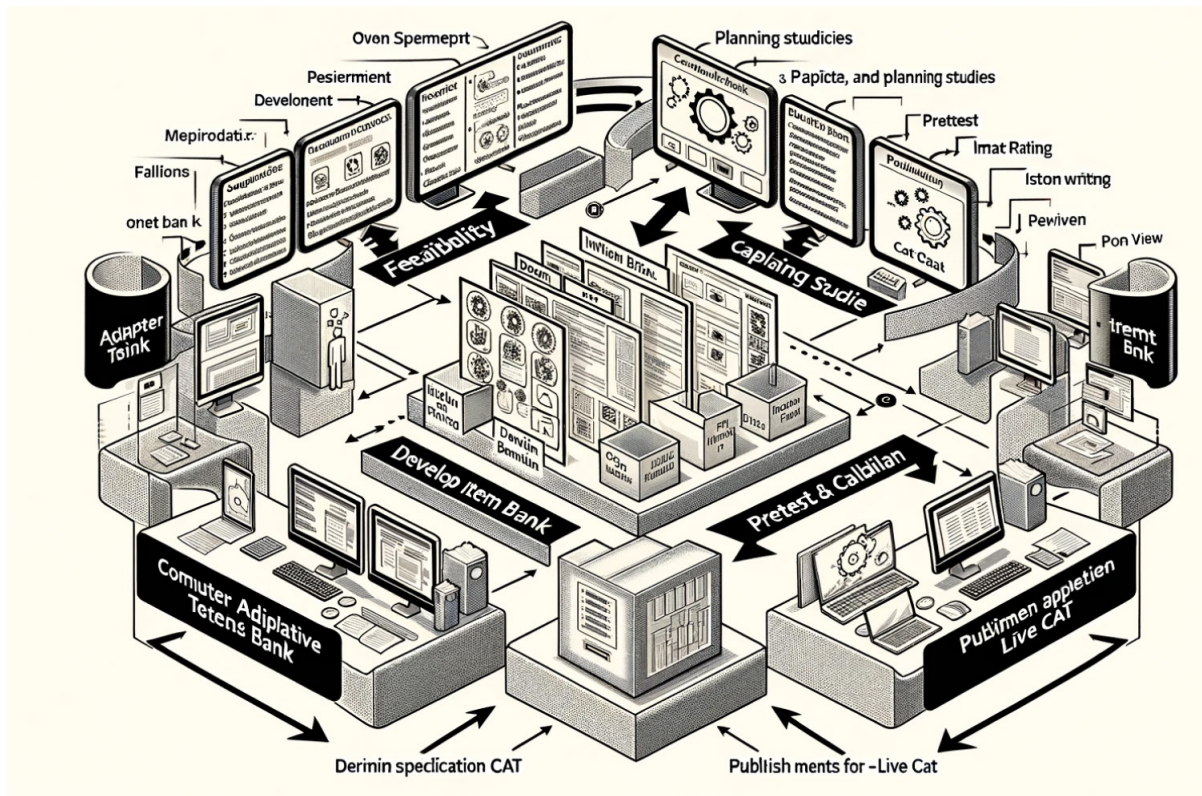
Implementing CAT in developing countries presents unique opportunities and challenges. In the context of Lesotho, the author is exploring the feasibility of ECoL adopting CAT to enhance the accuracy, fairness, and efficiency of its national assessments (see Figure 1 for the CAT framework). The transition to CAT in Lesotho could significantly improve the quality of educational evaluations by providing more precise measurements of student abilities and reducing the limitations associated with traditional testing methods. One of the primary opportunities presented by CAT in Lesotho is the potential to address educational disparities. Adaptive testing can tailor the difficulty of test items to match the individual abilities of students, thereby offering a more equitable assessment environment. This approach can help identify and support students who may be struggling, ensuring that interventions are timely

and targeted. Additionally, the immediate feedback provided by CAT can facilitate faster educational improvements, as educators can quickly identify and address learning gaps. The use of CAT in Lesotho also aligns with global educational trends towards more personalized and data-driven assessment methods. By adopting CAT, Lesotho can position itself at the forefront of educational innovation in Africa, setting a precedent for other developing countries to follow. The successful implementation of CAT in Lesotho could serve as a model for neighboring countries such as Eswatini, Zambia, Namibia, and so on demonstrating the feasibility and benefits of adaptive testing in similar contexts.

However, several challenges must be addressed to successfully implement CAT in Lesotho. The development of a robust technological infrastructure is essential, as CAT relies heavily on computer-based systems and reliable internet connectivity. Investments in hardware, software, and technical support are necessary to ensure the smooth operation of CAT systems. Additionally, digital literacy among educators, students, and administrators must be enhanced through comprehensive training programs to ensure that all stakeholders can effectively use the technology. The development of an extensive item bank is another critical component for the successful implementation of CAT. Items must be carefully calibrated using Item Response Theory (IRT) to ensure they accurately measure student abilities across various difficulty levels. This process requires significant expertise and resources, highlighting the need for collaboration with psychometricians and educational experts. Additionally, addressing concerns about test security and item exposure is crucial to maintain the integrity of the assessments.

By conducting a thorough feasibility study, including simulation studies using tools like SimulCAT, ECoL can identify potential challenges and develop strategies to overcome them. The insights gained from this research will be instrumental in guiding the successful implementation of CAT in Lesotho, ultimately leading to a more accurate, fair, and efficient assessment system. Simulation studies can help in fine-tuning the item selection algorithms, optimizing test lengths, and managing item exposure to ensure the CAT system is well-suited to the educational context of Lesotho. Furthermore, engaging stakeholders throughout the implementation process is crucial for the success of CAT in Lesotho. This includes involving educators, students, policymakers, and the community in discussions about the benefits and challenges of CAT. By fostering a collaborative approach, ECoL can build support for CAT and ensure that the system is responsive to the needs and concerns of all stakeholders. Such engagement can also help in identifying and addressing any logistical or cultural barriers that may arise during the implementation process.

Figure 1. CAT framework



7. FUTURE DIRECTIONS AND RECOMMENDATIONS FOR CAT IMPLEMENTATION IN LESOTHO

To ensure the successful implementation of CAT in Lesotho, several strategic steps and considerations must be addressed. Based on the findings from the feasibility study and simulation results, ECoL should develop a comprehensive roadmap that outlines the key phases of CAT implementation, including infrastructure development, item bank creation, stakeholder training, and continuous evaluation. Investing in the necessary technological infrastructure is crucial for the effective deployment of CAT. This includes securing reliable computer hardware, robust internet connectivity, and advanced software capable of administering and scoring adaptive tests. ECoL should collaborate with government agencies, private sector partners, and international organizations to secure funding and technical support for these initiatives. Additionally, establishing technical support teams to manage and troubleshoot the CAT systems will ensure smooth operation and minimize disruptions during assessments. To mitigate the challenges of limited technological infrastructure, ECoL could explore partnerships with

technology providers to facilitate access to affordable and reliable hardware and software solutions. Also, developing a high-quality item bank is essential for the success of CAT. This involves writing a large number of test items, calibrating them using Item Response Theory (IRT), and validating them through pilot testing with a representative sample of students. ECoL should engage psychometricians and subject matter experts to ensure the items are accurately calibrated and cover the full range of the curriculum. Regular updates and expansions of the item bank will be necessary to maintain its relevance and effectiveness over time. Additionally, ECoL should implement robust item security measures to prevent item overexposure and maintain the integrity of the item bank.

Training educators, administrators, and students on the use of CAT is vital for its acceptance and effective utilization. ECoL should develop comprehensive training programs that cover the technical aspects of CAT, its benefits, and its operational procedures. Workshops, seminars, and hands-on training sessions can help build digital literacy and ensure all stakeholders are comfortable with the new testing methodology. Additionally, engaging stakeholders in the planning and implementation process will foster a sense of ownership and support for CAT. ECoL should also consider creating online resources and support forums to provide ongoing assistance and address any issues that arise during the implementation process. Importantly, the implementation of CAT should be accompanied by continuous evaluation to monitor its effectiveness and address any emerging issues. ECoL should establish a system for collecting and analyzing data on test performance, user satisfaction, and technical reliability. Regular feedback from educators, students, and administrators will provide valuable insights into the strengths and weaknesses of the CAT system, guiding ongoing improvements and refinements. By leveraging data analytics, ECoL can identify trends and patterns that inform the continuous development of the CAT system and ensure it meets the evolving needs of the education sector.

Before fully implementing CAT across all national assessments, ECoL should conduct pilot tests in a controlled environment. These pilot tests will help identify potential challenges and allow for adjustments to be made before broader deployment. Based on the pilot results, ECoL can develop a phased rollout plan that gradually expands the use of CAT while ensuring that necessary support systems are in place at each stage. The phased implementation approach will allow ECoL to manage resources effectively and build capacity incrementally, ensuring a sustainable transition to CAT. Further, establishing a supportive policy and regulatory framework is essential for the successful implementation of CAT. ECoL should work with policymakers to develop regulations that support the use of CAT and address issues such as data privacy, test security, and fairness. Additionally, policies should be put in place to

ensure that CAT is accessible to all students, including those with disabilities or limited access to technology. By creating an enabling environment, ECoL can facilitate the smooth integration of CAT into the national assessment system. By following these recommendations, ECoL can successfully implement CAT, enhancing the accuracy, fairness, and efficiency of its national assessments. This transition will not only benefit examinees by providing more personalized and precise measurements of their abilities but also support educators in making data-driven decisions to improve educational outcomes in Lesotho.

8. CONCLUSION

The exploration of CAT for high-stakes assessment through a feasibility study by ECoL highlights both the potential benefits and the challenges associated with its implementation. CAT offers significant advantages in terms of precision, efficiency, fairness, and security, making it a promising alternative to traditional fixed-form assessments. However, successful implementation requires careful consideration of technological infrastructure, item bank development, stakeholder training, and continuous evaluation. The global adoption of CAT, driven by advancements in psychometric research and technology, demonstrates its effectiveness and adaptability across various educational contexts. For Lesotho, the implementation of CAT presents an opportunity to address educational disparities and enhance the quality of assessments, positioning the country at the forefront of educational innovation in Africa. By following a strategic roadmap that includes pilot testing, stakeholder engagement, and the establishment of a supportive policy framework, ECoL can successfully integrate CAT into its national assessment system, ultimately benefiting students, educators, and the broader educational community in Lesotho.

REFERENCES

- Ayanwale, M. A., Chere-Masopha, J., & Morena, M. C. (2022). The Classical Test or Item Response Measurement Theory: The Status of the Framework at the Examination Council of Lesotho. *International Journal of Learning, Teaching and Educational Research*, 21(8), 384-406. <https://doi.org/10.26803/IJLTER.21.8.22>
- Ayanwale, M. A., & Ndlovu, M. (2022). Transition from computer-based testing of national benchmark tests to adaptive testing: Robust application of fourth industrial revolution tools. *Cypriot Journal of Educational Sciences*, 17(9), 3327-3343. <https://doi.org/10.18844/CJES.V17I9.7124>

- Han, K. C. T. (2018a). Components of the item selection algorithm in computerized adaptive testing. *Journal of Educational Evaluation for Health Professions*, 15, 7. <https://doi.org/10.3352/JEE-HP.2018.15.7>
- Han, K. C. T. (2018b). Conducting simulation studies for computerized adaptive testing using SimulCAT: an instructional piece. *Journal of Educational Evaluation for Health Professions*, 15, 20. <https://doi.org/10.3352/jeehp.2018.15.20>
- Han, K. T. (2012). SimulCAT: Windows Software for Simulating Computerized Adaptive Test Administration. *Applied Psychological Measurement*, 36(1), 64-66. <https://doi.org/10.1177/01466216114144407>
- Han, K. T. (2016). Maximum Likelihood Score Estimation Method With Fences for Short-Length Tests and Computerized Adaptive Tests. *Applied Psychological Measurement*, 40(4), 289-301. <https://doi.org/10.1177/0146621616631317>
- Leroux, A. J., Lopez, M., Hembry, I., & Dodd, B. G. (2013). A Comparison of Exposure Control Procedures in CATs Using the 3PL Model. *Educational and Psychological Measurement*, 73(5), 857-874. <https://doi.org/10.1177/0013164413486802>
- Meijer, R. R., & Nering, M. L. (n.d.). *Computerized Adaptive Testing: Overview and Introduction*.
- Mujtaba, D. F., & Mahapatra, N. R. (2020). Artificial Intelligence in Computerized Adaptive Testing. *Proceedings - 2020 International Conference on Computational Science and Computational Intelligence, CSCI 2020*, 649-654. <https://doi.org/10.1109/CSCI51800.2020.00116>
- Ogunjimi, M. O., Ayanwale, M. A., Oladele, J. I., Daramola, D. S., Jimoh, I. M., & Owolabi, H. O. (2021). Simulated evidence of computer adaptive test length: Implications for high stakes assessment in Nigeria. *Journal of Higher Education Theory and Practice*, 21(2), 202-212. <https://doi.org/10.33423/JHETP.V21I2.4129>
- Oladele, J. I., Ndlovu, M., & Spangenberg, E. D. (2022). Simulated computer adaptive testing method choices for ability estimation with empirical evidence. *International Journal of Evaluation and Research in Education*, 11(3), 1392-1399. <https://doi.org/10.11591/IJERE.V11I3.21986>
- Oladele, J.I.; Ayanwale, M.A.; Owolabi, H.O. (2020). Paradigm Shifts in Computer Adaptive Testing in Nigeria in Terms of Simulated Evidences. *Journal of Social Science*, 63(1-3), 9-20. <https://doi.org/10.31901/24566756.2020/63.1-3.2264>
- Rice, N., Pêgo, J. M., Collares, C. F., Kisielewska, J., & Gale, T. (2022). The development and implementation of a computer adaptive progress test across European countries. *Computers and Education: Artificial Intelligence*, 3. <https://doi.org/10.1016/j.caeai.2022.100083>

- Seo, D. G. (2017). Overview and current management of computerized adaptive testing in licensing/certification examinations. *Journal of Educational Evaluation for Health Professions*, 14, 17. <https://doi.org/10.3352/JEEHP.2017.14.17>
- Stepanek, L., & Martinkova, P. (2020). Feasibility of computerized adaptive testing evaluated by Monte-Carlo and post-hoc simulations. *Proceedings of the 2020 Federated Conference on Computer Science and Information Systems, FedCSIS 2020*, 359-367. <https://doi.org/10.15439/2020F197>
- Thompson, N. (2023). Computerized Adaptive Testing (CAT): An Introduction. <https://assess.com/computerized-adaptive-testing/> Access date: 11-04-2024
- Thompson, N. A. (2009). Item selection in computerized classification testing. *Educational and Psychological Measurement*, 69(5), 778-793. <https://doi.org/10.1177/0013164408324460>
- Thompson, N. A., & Weiss, D. J. (2009). Computerized and Adaptive Testing in Educational Assessment. In F. Scheuermann & J. Björnsson (Eds.), *The Transition to Computer-Based Assessment: New Approaches to Skills Assessment and Implications for Large-scale Testing* (pp. 127-133). European Commission.
- Thompson, N. A., & Weiss, D. J. (2011). A framework for the development of computerized adaptive tests. *Practical Assessment, Research and Evaluation*, 16(1), 1-9.
- Tsaousis, I., Sideridis, G. D., & AlGhamdi, H. M. (2021). Evaluating a Computerized Adaptive Testing Version of a Cognitive Ability Test Using a Simulation Study. *Journal of Psychoeducational Assessment*, 39(8), 954-968. <https://doi.org/10.1177/07342829211027753>
- Viswanandhne, S., & Nandakumar, G. S. (2017). *Computer Based Adaptive Testing*. 2(2), 686-691.
- Wainer, H., Dorans, N. J., Flaugher, R., Green, B. F., & Mislevy, R. J. (2000). Computerized Adaptive Testing. *Computerized Adaptive Testing*. <https://doi.org/10.4324/9781410605931/Computerized-Adaptive-Testing-Howard-Wainer-Neil-Dorans-Ronald-Flaugher-Bert-Green-Robert-Mislevy>