# ASSESSMENT AND EVALUATION USING EVOLVING TECHNOLOGY

Omaze Anthony Afemikhe, MNAE, FASSEREN, KSM
Institute of Education, University of Benin, Benin City, Nigeria

## Introduction

Teaching and learning are two interrelated concepts in education. Teaching is expected to facilitate learning. The effectiveness of teaching and how much learning has been acquired by the students is important if the goals and objectives of education are to be realized. Obtaining information relevant in this regard requires that assessment must be conducted. Assessment may be seen as a product or as a process in that we can see the collection of requisite information as assessment and at the same time see the information collected as assessments. Collecting assessment information is not an end in itself; the information has to be used for decision-making about the learners, schools, policies, curriculum among others. This decision-making process which depends on an implicit or explicit standard is normally referred to as evaluation. The concept of evaluation as been seen as being elastic and therefore can be stretched to cover judgements of many kinds.

Technology is currently a major player in all facets of human life. It involves the use of scientific knowledge for practical purposes and encompasses a wide range of tools, machines, systems, and methods. It is used to solve problems, improve efficiency, and enhance human capabilities. Technological tools that easily come to mind include computers, smartphones, medical devices, and industrial machinery. As knowledge is not static but dynamic, technology keeps improving and discarding previously known ones. Thus, innovative advancements are currently developing or will be developed that they may impact society in a great way. Break throughs in artificial intelligence, biotechnology, nanotechnology, and renewable energy are examples of emerging technologies.

These emerging technologies are now being used in assessment and evaluation in education. How they have been used, the impact they have made in education and future directions are parts of the thrust of this presentation. The presentation will explore the following issues: Overview of traditional assessment methods and alternative assessment, current trends in technology-enhanced assessment, benefits and challenges of evolving technology in assessment and future directions.

**Overview of traditional assessment methods and alternative assessment**

The educational landscape as been dominated by traditional assessments for many years. When used what readily comes to mind are assessment tools which sometimes are standardized and focus on quantifiable achievements through tests, quizzes, and exams and include multiple-choice tests, true/false tests, short answers, and essays. These assessments are characterized by offering clear, objective measures of student performance against predetermined criteria, and their structured format. Formalized traditional assessment can be traced back to ancient civilizations, with one of the earliest examples being the Chinese Imperial examination system. Its was used in identification of talent and provision of reward across imperial China. It was a rigorous written tests to select candidates for government positions, emphasizing the importance of standardized testing.

The use of oral assessments has its origin in antiquity and continues to be used in education. However, in the Western world, formal written assessments began to replace oral examinations in the 19th century. This was attendant to the expansion of the public education system which was aimed at educating the masses rather than only the elite. (Center on Education Policy, 2020) The development of standardized tests during this period was driven by the need for objective and reliable methods to evaluate student performance and ensure educational quality.

With advancement in statistical methods and psychometrics, traditional assessments developed with the use of multiple-choice tests, essays and other standardized formats. These methods have continued to be used as they provide consistent and comparable measures of student achievement. They facilitate comparisons across educational settings and can be administered to large groups simultaneously in addition to an ability to provide a uniform benchmark for measuring student achievement, hence are they time and resource efficient. However, they often fail to capture the totality of the breadth of student learning and may not capture assessment of critical thinking or problem solving. They could also be prone to stress and anxiety because of the high stakes always associated with them.

To mitigate some of the problems associated with traditional assessments, there has been a paradigm shift towards more inclusive and holistic approaches to evaluating student learning. This shift has led to assessments which are seen as authentic which unlike their traditional counterparts are not solely reliant on standardized tests. They include portfolios, presentations, peer- assessment, group projects, and self-assessments. They aim to provide a more comprehensive view of a student's abilities and progress. Alternative assessments provide a broader perspective on student learning, they also capture critical thinking, creativity, and collaboration skills in addition to academic achievement. As students are involved in these diverse and meaningful assessment tasks, these methods can boost motivation and student engagement.

There has also been a shift towards innovative test items. A comprehensive conception of an innovative test according to Parshall et al. (2000) goes beyond the "discrete, text-based, multiple-choice format". Such items have simple enhancements which could include the addition of 'a video clip to a multiple-choice item stem, new ways of interacting with the stimulus (e.g., sortable tables or searchable documents) to complex open-ended response formats such as

performance simulations or speech capture' (Brunner & Becker, undated). Traditional select items have been criticized for not being authentic and this criticism makes innovative items to hold a lot of sway. Innovative items have greater fidelity to the real world as they can offer greater *construct representation.* According to Sireci and Zenisky (2006, p. 300), construct representation is the "the ability of a test to fully represent all the knowledge, skills, and abilities inherent in the construct measured".

Conventional tests have been criticized as imperfect measures of constructs as they can leave out what should be included or include that which should be omitted or both. While innovative item types can possibly tackle the 'leaving something out' or construct representation concern, they however stand the risk of introducing measurement error by including 'something that should be left out' thus leading to the presence of *construct-irrelevant variance*. Authenticity and directedness have been seen as two concepts that have a bearing to threats to validity to underrepresentation and construct irrelevant variance respectively. Consequently, the need to explore the convergent and divergent validity becomes very relevant in the construction of innovative items.

Although there are several ways to categorize innovative items, one taxonomy provided by Parshall et al. (2000) is sufficiently comprehensive and provides five dimensions: item format, response action, media inclusion, level of interactivity, and scoring method, or algorithm. However, with development in technology, the taxonomy has been enlarged to include assessment applications that are increasingly divergent from the traditional testing environment. The revised taxonomy has been expanded to a movement away "item format" into the broader term "assessment structure," in addition to the inclusion of the new dimensions of complexity and fidelity. The seven levels of the revised taxonomy are: assessment structure, complexity, fidelity, interactivity, response action, media

inclusion, and scoring algorithm. The move into innovative assessments has been facilitated using computers and associated technology. This is one way technology has come to enhance assessment and evaluation.

**Assessment and evaluation terrain**

Discussions so far may tend to indicate that interest in assessment and evaluation is on tests and their construction. Some aspects include item metadata, Test theories, item fairness and equity, measurement invariance, item banking and equating. It is important to review these components before attempting to look at how evolving technologies can be applied comprehensively in assessment and evaluation.

**Measurement theories**

Two measurement theories have dominated assessment scene, and they are Classical Test Theory (CTT) and Item Response Theory (IRT). In CTT the observed score is seen as a combination of two things: the true score and random error. This is represented as:

$$X_{observed} = X_{True} + X_{error}$$

The true score is your actual ability or trait level, and the error is all the random stuff that could affect the score (like if you were tired or distracted during the test) CTT assumes that errors are random and will cancel out across large numbers of test-takers. There are a few key assumptions in Classical Test Theory:

1. **Linearity**: The observed score is the sum of the true score and the error score.
2. **Random Error**: Errors are random and average out to zero across a large number of test-takers.
3. **Consistency**: The true score remains stable across multiple test administrations.
4. **Equal Error Variance**: The error variance is the same for all test-takers.

These assumptions make CTT a straightforward and useful tool, but when the assumptions do not hold:

1. **Error Variance**: If the assumption of equal error variance fails, the error might not
   average out across individuals, making the results less reliable.
2. **True Score Instability**: If the true score isn't stable, multiple test adminis trations
   might yield inconsistent results.
3. **Non-Random errors:** If errors are not random, they might systematically bias the results.
4. **Item Difficulty**: CTT doesn't account for the varying difficulty of individual test items, which can affect the validity of the test scores.

With these limitations, there can be inaccuracies and misinterpretations of test results. That is why more advanced methods like Item Response Theory (IRT) are sometimes preferred, as they address some of these issues more effectively.

Item Response Theory (IRT) has the following core assumptions:

1. **Unidimensionality**: Assumes a single latent trait (e.g., ability) influences responses to test items.
2. **Local Independence**: Given a person's trait level, their responses to individual items are independent of each other.
3. **Monotonicity**: Higher levels of the latent trait should increase the probability of a correct response.

These assumptions allow IRT to provide detailed item-level information, making it more nuance compared to CTT. Different parameter models (1, 2, 3 and 4) have been used in the literature.

**Item metadata**

These are characteristics of test items which include Difficulty Level, Discrimination Index, Content Area, and Cognitive Skill Assessed. These

characteristics particularly Difficulty level and discrimination index can be computed using either the CTT or IRT.

**Item Fairness and Equity**

It involves the analysis and identification of any biases in test items that might disadvantage certain groups of test-takers based on Cultural Background, Gender, Socioeconomic Status, etc

**Measurement invariance**

This focuses on measurement consistency, and it is the property that ensures test items measure the same construct across different groups of test-takers, ensuring Cross-Group Comparability, Equitable Assessment Outcomes, Validity Across Diverse Populations

**Item banking**

Item banking is a process of developing and managing a large pool of high-quality forms of assessments, such as tests, quizzes, or examinations. test items which are usually stored in a central repository called an item bank (Oguche and Oguche, 2015). The intent is to improve the efficiency and quality of test development by having a ready-to-use collection of questions that can be drawn upon to generate various test forms. An item bank can be generated using the following steps: define the purpose and scope, develop item specification, write items, review and edit items, pilot test to a small group, revise and refine, store and organize, and maintain and update.

**Equating:**

When two or more test forms have been constructed a common measurement scale is desired to compare scores across them (Dorans et al., 2010). The statistical control for difficulty differences so that scores can be used interchangeably across forms is the main objective of this procedure. Equating is necessary if different forms are used across different administrations to ensure

fairness and consistency. There are different types of equating such as linear equating, equipercentile equating, and mean equating.

These are some issues of interest in assessment and evaluation. We will look at which evolving technologies can be used to handle them in this presentation. Before then we will look at the scope of emerging technology.

**Emerging Technology-**

Technology use has come to stay. Whatever technology is applied normally passes through some stages the totality of which have come to be known as technology cycle with the following stages:

- ✓ Innovation which is the development and deployment of the technology:
- ✓ Growth involving rapid adoption and improvement of the technology.
- ✓ Maturity is the widespread use and optimization of the technology.
- ✓ Decline stage is a gradual obsolescence as newer technologies emerge.

Technology has a long history dating back to when the first tools were made by man. It progressed through the Neolithic Revolution when several key technologies arose together. Humans moved from getting their food by foraging to getting it through agriculture. Steam engine, boat and telegraph, telephone and combustion engine followed in that order. In 1937, the computer was designed by John Atanasoff. Personal computer and internet followed thereafter. The introduction of the internet and the personal computer have greatly revolutionized technology use in many facets of life including education. Technology continues to grow, and hence new variants are emerging hence the name emerging technologies.

These are new and innovative technologies that are being developed or have recently been introduced into the market. They have the potential to significantly impact all segments of society. They can be recognized by their common

characteristics which include high potential, uncertainty, rapid evolving, interdisciplinary, and destructive character. Evolving technologies refer to innovations that are continuously developing and advancing, often transforming industries and daily life. These technologies typically start as emerging concepts and gradually become integral parts of our lives as they mature and improve. Examples include key examples and their evolution: Artificial Intelligence (AI), Renewable Energy Technologies, Biotechnology, Internet of Things (IoT), 3D Printing, Blockchain, Augmented Reality (AR) and Virtual Reality (VR) and Mixed reality (MR), among others.

The use of these technologies is ubiquitous. Artificial Intelligence is utilised in decision-making and automation processes in healthcare, finance, and transportation to improve decision-making and automate processes. Internet of Things (IoT) is used in manufacturing, logistics, and agriculture to monitor and control processes and optimize resource usage. Blockchain is used in finance, supply chain management, and healthcare to provide secure and transparent transactions. AR and VR are being used in education, entertainment, and healthcare to enhance the user experience and provide immersive environments. Our interest is the use of these technologies in education with particular focus on assessment and evaluation.

**Current Trends in Technology-Enhanced Assessment**

Technology-enhanced assessments are transforming the way educators evaluate and support student learning. Technology makes possible more flexible, tailored presentations to students of a much wider and richer array of tasks and environments where students can learn and where they can show us what they know and how they know it. Thus, there ‘is an interesting and powerful confluence among theory, research, technology, and practice, especially when it comes to the integration of curriculum, instruction, and assessment’ (Pellegrino & Quellmalz, 2010). Here are some current trends in this area:

**Adaptive Learning and Assessment**

Adaptive assessments use algorithms to adjust the difficulty of questions based on the student's previous answers. When implemented in the classroom it can significantly enhance personalized instruction and improve student outcomes. The implementation can be executed with the following steps:

**1. Identify Your Goals**

What of interest to the teacher can vary dep This is a provider that personalizes educational content to meet individual student needs. The interest could be aiming to improve student engagement, tailor instruction to individual needs, or enhance assessment accuracy? The exact goal has to be clearly understood and all decisions in use of adaptive assessment hinge on the goal identified.

**2. Choose the Right Technology**

What tools or platforms are selected must align with the goals. There are many platforms that can be used. Some adaptive learning platforms include

- ✓ Knewton: It uses data-driven algorithm to create optimized learning paths, ensuring that each student receives tailored instruction based on their performance and learning style. Knewton adaptive learning combines a data-driven approach to course materials with a 360° view of an individual student to create optimized learning paths that help all learners reach their highest potential. It provides real time feedback. The underlying principle is the belief that each student is unique and that their knowledge is not defined by their last answer. Sometimes what a student needs may not be the next lesson in their textbook. Alta considers thousands of data-points through continuous measurement of the content's instructional value in conjunction with ongoing assessment of student learning proficiencies on every concept studied. This creates unique, responsive learning paths for students.
- ✓ DreamBox Learning: This platform offers adaptive math and reading programmes designed to personalize instruction for students from

kindergarten through 12th grade. The mathematic component focuses on developing conceptual understanding and critical skills while the reading component is designed to improve fluency, comprehension and vocabulary with data-driven personalized instruction.

- ✓ Smart Sparrow: it is a learning design platform that enables educators to create interactive and adaptive online learning experiences. The 'platform allows you to easily create visually rich online courseware. Decide your lesson structure and the platform enables dragging and dropping elements like images and videos, or even allows import of interactive components, picking a theme and screen template…' It is a powerful authoring tool which makes it easy to create impactful and stunning experiences and shape the way the courseware adapts to the needs of each learner. It has a detailed analytics dashboard, and the reports provide actionable insights, beyond grades. With its use one can choose from hundreds of widgets, simulations, and games to engage your learners. (https://www.smartsparrow.com)

**3. Integrate with Your Curriculum**

Ensure that the adaptive learning tools you choose can be seamlessly integrated into your existing curriculum. This might involve aligning the content and assessments with your learning objectives and standards.

**4. Train and Prepare**

The teacher needs to familiarize his/herself with the platform's features and functionalities. Training is then provided for the teachers and the students on how to use the adaptive learning tools effectively.

**5. Start Small**

As with all programmes, it is important to start with a pilot study or a single unit to test the effectiveness of adaptive learning in your classroom. This will help to identify any challenges and make necessary adjustments before a full-scale implementation.

**6. Monitor and Adjust**

Regularly monitor student progress and engagement with the adaptive learning tools. Use the data and insights provided by the platform to adjust your instruction and provide additional support where needed.

**7. Encourage Student Feedback**

Feedback is gathered from the students about their experiences with adaptive learning. This can help to understand what works well and what might need improvement.

**8. Collaborate with Colleagues**

Share your experiences and insights with other educators. This collaboration can lead to the exchange of best practices and innovative ideas for implementing adaptive learning.

**9. Evaluate and Reflect**

Periodically evaluate the impact of adaptive learning on student outcomes. Reflect on what has been successful and what could be improved and adjust accordingly.

**10. Stay Updated**

Keep up with the latest trends and advancements in adaptive learning technology. Continuous learning and adaptation help the teacher to make the most of these tools in your classroom.

**Digital assessments**

This is also seen as e-Assessment or electronic assessment is used to present assessments, tests, surveys, and other measures using computers, tablets, and mobile phones which are examples of digital devices. AI and ML play a very important role in this respect. With digital assessment, one is able to develop items, publish tests, deliver tests, and provide meaningful results in quick way and also validly. Digital assessment platforms that are cloud-based are required to deliver digital assessment. Such platforms do much more than just the delivery

though, as they are also used in Item banking which involves writing, reviewing, and determining metadata, as well as online test delivery, remote proctoring, adaptive testing and essay scoring.

There are many software that are used as part of digital assessment for different purposes. Among these are:

- ✓ Kahoot: This gamification platform helps teachers build the learning process in a form of a game by creating multiple choice questions or using already existing games.
- ✓ Socrative: This software is equipped with various features that can be used by educators for many different purposes. It also includes activities like quizzes, quick questions, exit tickets, etc.
- ✓ Plickers is another real-time educational software allowing to get data for an assessment. It is a comprehensive tool that helps teachers understand exactly where their students are in terms of progress.
- ✓ Google Forms allows teachers to create multiple choice questions in the form of a survey and enhance it with images and videos in a few minutes.

The learner is now more than before seen to be at the centre of teaching and learning. Technology provides an opportunity for assessment to be integrated into the learning process with corresponding adjustments. Furthermore, with digital technology there is a shift in focus from mere content to critical thinking. The advantages of this integration of technology according to Issayeva (2024) include **accessibility** as there is ease of use for staff and learners. Questionnaires can easily be setup, grading methods determined, and invitations sent to examinees. Examinees can take assessments remotely. There is **transparency** as educators can evaluate performance of a group against an individual learner for analytical and pedagogical reasons. It can equally generate reports as so educators are able to identify learning problem areas on both individual and group levels soon after

assessments occur to adapt to learners' needs, strengths, and weaknesses. Learners are also provided with instant feedback, unlike traditional paper exams.

There is also the benefit of **profitability** because of flexibility of venue and time, reduced human, logistic and administrative costs lend considerable pre-eminence to electronic assessment over traditional exam settings. **It is eco-friendly** as institutions can go paper-free and avoid printing exam papers and other materials. The storage space is taken care of as all data can be stored on a single server, especially in respect to keeping records in paper. There is enhanced privacy with chances of cheating highly reduced. **Enhanced privacy** for students is another advantage of digital assessment that validates its utility. There is a tiny probability of malicious activities, such as cheating and other unlawful practices that can potentially rig the system and lead to incorrect results. An **automated grading system** which is more convenient and time-efficient than standard marking and grading procedures that minimizes human error is provided. Automated scoring juxtaposes examinees' responses against model answers and makes relevant judgements. Automated item generation, item banking, test assembling and publishing saves precious time that would otherwise be wasted on repetitive tasks.

The challenges include difficulty with grading long-answer questions. Technology use in the education sector continues to evolve and even essays can already be marked digitally with a help of AI-features on some platforms like ***FastTest***. Adaptation will be required as resistance may be noticed and there could be need to upgrade the system. As in developing counties like Nigeria, technology is not always reliable, and some locations cannot provide all examinees with stable access to electricity, internet connection, and other basic system requirements. This can be problematic. It could also be very costly to implement due to use of wrong strategies as part of the planning stage.

## Gamification

Gamification has its root in the word game. It involves the application of game-design elements and mechanics to non-game contexts, such as learning, work, or marketing. Elements like timing, points scoring, badges, leaderboards, and challenges are normally incorporated to make activities more engaging and motivating. It involves making non-game activities feel like they're games. It's a way of adding extrinsic motivation. According to Chou (2015) to successfully gamify an activity, it needs to tap into one of eight "core drives" of human motivation: Meaning, Purpose, Autonomy, Mastery, Social Status, Relatedness, Scarcity, and Loss Aversion.

There are different strategies involved in gamification which include discussion boards, quizzes, jeopardy, classroom response system, game-based learning systems and game-enhanced learning system. While discussion boards and quizzes are on-line gamification strategies, jeopardy and classroom response are in-class strategies and game-based learning systems, and game-enhance learning system are out-of-class strategies. For example, one can require students to make comments on an issue with comments earning points for contributions or replies. For any of such comments or replies we can award experience points (XP) that are a common element in learning boards and online platforms. They provide a sense of achievement and progress, encouraging one to continue learning. consider creating a narrative or quest that draws learners in and helps them see the consequence of their responses. In quizzes, interactive narrative could be provided with each question leading into the next, and may build upon previous answers, all the while being part of a larger narrative or story that compels the learner to remain engaged. Such narratives may even include hints that the learner can choose to use or not. Whenever gamification is used there is need to apply

rules, goals, interaction, feedback, problem solving, competition, story, and fun (Vandercruysse et al., 2012).

**Artificial Intelligence and Machine Learning**

Artificial Intelligence (AI**)** is a branch of computer science that aims to create intelligent agents, which are systems that can reason, learn, and act autonomously. The purpose of the agents is to mimic human cognitive functions, such as understanding natural language, recognizing patterns, and solving problems. Through AI, a computer system uses mathematics and logic to simulate the reasoning that people use to learn from new information and make decisions. Machine Learning (ML) is a subset of artificial intelligence that focuses on algorithms that allow computers to learn from data and improve their performance over time. It involves training computers to recognize patterns, make predictions, and solve problems without being explicitly programmed.

AI and ML are revolutionizing the way assessments and evaluations are conducted. They offer significant advantages over traditional methods, such as increased efficiency, objectivity, and personalized feedback. Some key applications of these technologies in assessment and evaluation are automated grading**,** adaptive assessment, item analysis, plagiarism detection and predictive analysis. In automated grading, AI can be used to score essay as it can analyze essays for grammar, coherence, and content, providing more consistent and objective scoring than human graders. ML algorithms are used for grading multiple-choice test items quickly and accurately, freeing up teachers' time for more personalized instruction. As part of adaptive assessment, as mentioned earlier AI can tailor assessments to each student's individual needs and abilities, adjusting the difficulty level of questions in real-time, while ML algorithms can recommend personalized learning paths based on student performance, ensuring that learners are challenged and supported appropriately.

The quality of test items is an important consideration in assessment. Issues that are germane include item characteristics such as item difficulty, item bias among others. AI can analyze test data to identify questions that are too difficult or too easy, helping teachers improve the quality of their assessments. ML can detect bias in test items, ensuring that assessments are fair and equitable for all students. AI can automate the process of conducting multigroup Confirmatory Factor Analyses to compare the factor structure and loadings across different groups. Bayesian methods can provide more flexible and robust estimates of measurement invariance, especially when dealing with small sample sizes or complex models. ML algorithms can efficiently detect DIF, identifying items that function differently for different groups. AI can handle large datasets and complex models, making it suitable for detecting DIF in large-scale assessments. AI can fit IRT models to data and assess the extent of measurement invariance. ML can improve the accuracy and efficiency of IRT parameter estimation, leading to more reliable results. Combining multiple ML models can improve the accuracy and robustness of measurement invariance detection. Deep learning architectures can capture complex patterns in data, potentially improving the detection of subtle forms of measurement invariance.

Furthermore, AI-powered plagiarism detection tools can compare student work to a vast database of sources, identifying instances of plagiarism. And ML can analyze student data to predict their likelihood of success in a course or program, allowing for early intervention and support. Personalized feedback is also possible as AI can generate personalized explanations for concepts, considering the learner's prior knowledge and learning style. This can help learners better understand complex topics and make connections between different ideas. It also can provide immediate feedback on learner performance, allowing them to identify areas where they need to improve and adjust their learning strategies. In

addition, AI can recommend specific resources, such as articles, videos, or quizzes, that are relevant to the learner's individual needs and interests. Personalized feedback on assignments is also possible with AI taking into consideration the specific strengths and weaknesses of the learner. This can help learners identify areas for improvement and develop their writing or problem-solving skills.

**Benefits and challenges in use of Evolving Technology in Assessment and evaluation**

Some of the benefits of evolving technologies may have been tangentially alluded to in this presentation. In this section an attempt is made to show how relevant they are in assessment and evaluation in education. The benefits can be categorized into economic, social and pure educational benefits. Economically, new technologies often streamline processes leading to higher productivity and reduced cost. This increased efficiency is noticeable when the evolving technologies are used in educational assessment and evaluation. Time is saved in executing the many processes required to ensure that assessments are doing what they should be doing. In education, teachers may not all be proficient in the use of the evolving technologies and therefore opportunities are opened in term of job creation to support the work of teachers in utilization of the technologies in delivering on their mandate. In the same vein, they can drive educational growth and development.

Socially, they can enhance quality of life as they provide solutions to challenges in education. This is noticed in adaptive assessment, automated assessment as many time-consuming tasks can easily be handled thereby freeing up teachers and administrators to focus on other important aspects of their work.ML can as stated earlier can use assessment information in recommending personalized learning paths based on student performance, ensuring that learners are

challenged and supported appropriately. There is also increased connectivity consequent on advancements in communication technologies as people worldwide are connected, fostering collaboration among assessment experts. People with disabilities can easily be provided with services and information.

Educationally, new technologies can personalize learning, making it more engaging and effective. Online learning platforms and digital resources can expand access to education, especially in remote areas and these can equally be easily assessed. They are also useful in assisting individuals to develop essential skills for the future, such as critical thinking, problem-solving, and digital literacy.

Despite these benefits, there are challenges that have to be addressed to reap full benefits from the use of evolving technologies. Among these are technological challenges that include the complexity to understand and implement them as they require specialized knowledge and skills. Sometimes different technologies that coexist may not be compatible and this can be a challenge. Data is at the core of technology use and these have to be protected from cyber threats. We may not be sure of the reliability of new technologies, and this can lead to potential disruptions or failures. Ethically, issues of privacy come up when we collect and use personal data, especially with the rise of data-driven technologies. When there are biases in the data, Algorithms and AI systems can perpetuate them based on their training, thereby leading to unfair outcomes. Variability in access to technology can exacerbate existing social and economic disparities (Rejo, 2021).

Economically, acquisition and implementing new technologies can be expensive as significant investments may be required. Very often we need to calculate the economic benefits and in education this could be difficult because of its social nature. Furthermore, the dynamic nature of technologies makes them to go

obsolete easily and therefore constant replacement involving cost becomes a concern.

These challenges can affect how well emerging technologies are used and there is need to attempt how to mitigate them. First there is need to develop ethical guidelines and regulations for the use of technology in assessment and evaluation. Educators and educational institutions should be equipped with the skills needed to leverage new technologies. This may look impracticable but a support specialist available to assist institutions and individuals could be one way out of it. Furthermore, there is need to evaluate the potential social, economic, and environmental consequences of new technologies before adopting them for use.

**Future Directions**

Technology will continue to advance rapidly, and the future of assessment and evaluation is poised for significant transformation. Statistics is a bed rock of assessment and evaluation. Presently, with online testing and online learning behaviours, large-scale data has now become available to engross educational statistics. To move educational testing statistics into the dynamic analysis of big data with machine learning and artificial intelligence algorithms computational psychometrics has come to the rescue. This potentially will create a black box of sophisticated statistical models for learners, teachers, administrators, and citizens. The introduction of computing technologies will further strengthen automation of item generation (Gierl & Lai, 2016) and scoring of performances (Shin et al., 2021). As indicated by Linden and Glas (2000) customisation of test content according to test-taker performance is thus possible. The introduction of online and distance testing has become a commonplace practice following the COVID-19. The consequence is that concerns about how technology is used to assure the integrity of student performance (Dawson, 2021).

With respect to personalized adaptive learning Brown (2022) opine that assessments will become increasingly personalized, adapting to each learner's individual pace, strengths, and weaknesses in real-time. AI-powered systems will provide tailored feedback, explanations, and guidance, simulating one-on-one tutoring experiences. Virtual reality (VR) and augmented reality (AR) are two different types of technologies. While Virtual reality (VR) is a computer-generated simulation of a real-life environment that can be explored and interacted with using head-mounted displays or other devices (Stuchlikova et al., 2017), Augmented reality (AR) involves superimposing of digital information on the user's real-world view through specialized glasses, lenses, or contact lenses, or through an app that projects virtual images overtop of a live camera feed from a mobile device. This is done to augment what would appear to be experiential in the real world (Shanu et al., 2022). VR provides an immersive environment through which assessments will be delivered, allowing for more authentic and engaging evaluation of skills, such as problem-solving and decision-making. Both AR and VR would be used to create realistic scenarios to assess learners' abilities in practical courses.

A digital record book that uses cryptography to securely organize and store data from ledger entries is a blockchain. Thus, a blockchain credential is when a business or person's background qualification is uploaded to a block on a digital ledger for improved security, verification, and enhanced accessibility. This makes blockchain a veritable tool to secure and allow records to be verified. It allows for authentic verification of credentials thereby reducing fraud and enhancing trust. Neurotechnology-Enhanced Assessments using Brain-computer interfaces (BCIs) can measure brain activity during assessments, providing insights into cognitive processes and learning styles. As a result, it will be possible to use it to adjust the difficulty of assessments based on real-time brain activity.

AI algorithms will be developed to be fair and unbiased, ensuring that assessments do not perpetuate discrimination. Furthermore, AI systems will become more transparent, allowing users to understand how decisions are made. With respect to continuous assessment and micro-credentials, it will be possible to assess smaller units of learning continuously, thereby allowing learners to accumulate micro-credentials that demonstrate specific skills and competencies. Immediate feedback, that will enable learners to adjust their learning strategies as needed will be provided by continuous assessment. Assessments will incorporate game-like elements to make learning more enjoyable and motivating through gamification and stimulation. Immersive learning experiences will provide opportunities for learners to practice skills in realistic scenarios.

**Conclusion**

Emerging technologies, such as AI, machine learning, and adaptive testing platf orms, are revolutionizing traditional assessment and evaluation methods. These technologies offer increased efficiency, precision, and the ability to tailor assess ments to individual needs, thereby enhancing the overall learning experience. However, they also bring new challenges, such as ensuring data privacy, maintai ning equity, and addressing the potential for bias in automated systems. There is however need to balance innovation with ethical considerations maximize the benefits while mitigating risks. In short, while emerging technologies hold great promise for the future of assessments, their integration must be handled thoughtfully to truly enhance educational outcomes.

**References**

Brown, G.T.L. (2022) The past, present and future of educational assessment: A transdisciplinary perspective. Frontiers in Education 7:1060633. doi: 10.3389/feduc.2022.1060633

Center on Education Policy (2020), History and Evolution of Public Education in

the US. https:// files.eric.ed.gov/fulltext/ED606970.pdf

Chou, Y. (2015). Actionable Gamification: Beyond Points, Badges, and Leaderboards. McGraw-Hill Education

Dawson, P. (2021). Defending Assessment Security in a Digital World: Preventing e-cheating and Supporting Academic Integrity in Higher Education. Routledge.

Dorrans, N.J., Moses, T.P. and Daniel R. Eignor, D.R. (2010). Principles and practices of test score equating. Educational Testing Service (ETS)1.

Gierl, M. J., and Lai, H. (2016). A process for reviewing and evaluating generated test items. Educational Measurement Issues Practice 35, 6–20. doi: 10.1111/emip.12129.

Issayev, L. (2024). What is Digital Assessment, aka e-Assessment? https://assess.com/What is digital assessment.

Knewton. https://Support.knewton.com/s/article/Knewton Adaptive Learning and How it's Unique.

Linden, W. J. V. D., and Glas, G. A. W. (2000). Computerized adaptive testing: Theory and practice. Kluwer Academic Publishers.

Oguche, R. And Ogwuche, F.E. (2015). Item banking and test administration examining effectiveness in the UTME and NABTEB. Academic Scholarship Journal. Vol.10. 1, 147-159.

Parshall, C. G., Davey, T., & Pashley, P. J. (2000). Innovative item types for computerized testing. In W. J. van der Linden & G. A. W. Glas (Eds.), Computerized Adaptive Testing: Theory and Practice (pp. 129-148). Springer, Dordrecht https://doi.org/10.1007/0-306-47531 6_7

Parshall, C. G. & Harmes, J. C. (2008, March). Stages in designing innovative item types. Presented at the annual meeting of the Association of Test Publishers, Dallas, TX.

Pellegrino, J. W. & Quellmalz, E.S. (2010). Perspectives on the integration of

technology and assessment Journal of Research in Technology Education, Vol. 43, No. 2, pp. 119–134

Sireci, S. G. & Zenisky, A. L. (2006). Innovative item formats in computer-based testing: In pursuit of improved construct representation. In S. M. Downing & T. M. Haladyna (Eds.). Handbook of Test Development. Routledge Handbook of test development (pp. 329-347). Lawrence Erlbaum Associates.

Rejo, T. G. (2021) Issues of Assessment, Measurement, and Evaluation. SBS Journal of Applied Business Research (JABR). Vol VII, 15-30.

Shanu S., Narula D., Nayana, Pathak L. K., Mahato S. (2022). AR/VR Technology for autonomous vehicles and knowledge-based risk assessment. In virtual and augmented reality for automobile industry: Innovation vision and applications (pp. 87–109). Springer International Publishing

Shin, J., Guo, Q., and Gierl, M. J. (2021). "Automated essay scoring using deep learning algorithms," in Handbook of research on modern educational technologies, applications, and management. ed. D. B. A. M. Khosrow-Pour (ed). IGI Global.

Sparrow. https://www.smartsparrow.com

Utah Valley University (2024). Understanding Technology. Utah Valley University.

Vandercruysse, S., Vandewaetere, M., & Clarebout, G. (2012). Game-based learning: A review on the effectiveness of educational games. In M. M. Cruz-Cunha (Ed.), Handbook of research on serious games as educational, business, and research tools (pp. 628–647). IGI Global.