

Introduction of Smart Theory-based Examinations STBES in the Nigerian Education Policy

Nwosu Ogochukwu Fidelia¹, Nwosu Okwudili Chukwuma², Chilaka Akuchinyere Eunice¹,
Onyemachi Chinedu Okechukwu¹

¹Department of Computer and Robotics Education, University of Nigeria

²Department of Public Administration and Local Government, University of Nigeria

Corresponding Author: Chilaka Akuchinyere Eunice. akuchinyere.chilaka@unn.edu.ng

Abstract

In Nigeria education system, theory-based questions are marked manually by different people from different background, perceptive, approaches and in different moods at the time of marking. The subjectivity of scoring, halo effects, examiner mood, extraneous elements like handwriting legibility and ink color, and incorrectly summing up scores are some of the other factors that impact the grading of theory problems. Furthermore, in the event that the test is not assessed anonymously, the score may also be influenced by the grader's personal sentiments toward the student, the student's sex or race, or teacher expectations regarding the caliber of the student's work. Thus, a single exam that is externally given, graded, and has the potential to elevate or depress the student is all that stands between them and their destiny. Hence, in order to ameliorate the above mentioned problems, the study aims at testing the efficacy of a smart theory-based examinations system in the Nigerian education policy. The design for the study was Research and Development where automated system was developed for marking theory-based examination questions. The population of the study was 15 JSS 3 students that were used to test the validity and reliability of the developed system. It was discovered in the study that the STBES was effective and less time consuming and will help eliminate the possibility of guess work and lead to results that do not reflect the student's actual proficiency in the course. Formulation of policies by government for the introduction of STBES for marking theory based examinations by examination bodies in Nigeria was therefore recommended.

Keywords: Theory-Based Examination, Examination, Smart system, Nigerian Education Policy

Introduction

Since computers have the ability to enhance every system to which they are applied, they have favorably permeated and affected every aspect of human effort. This worldwide trend of automating systems includes the education sector as well. Teachers are required to create engaging and dynamic learning environments using the modern educational tools and methodologies of today. In Nigeria, certification is awarded to those who successfully finish a course of study with high academic standing, regardless of their educational level. Students are thus required to sit for public examinations administered by examination

boards upon completion of their secondary school studies. Therefore, in order to determine if a student deserves to be promoted to a different class, teachers rely on these accomplishment tests/examinations for a thorough assessment of their pupils.

According to Clark (2014), the Nigerian education system has a manual marking system for theory-based questions that is completed by individuals with varying backgrounds, perspectives, methods, and emotions. Subjectivity of scoring, or the fact that different teachers assign different scores to the same questions, is one of the factors influencing the grading of theory questions, which may bring about discrepancies in the

scores of two persons marking same paper at different point, even with same marking guide. This may be because of the mood at the time of marking, halo effects, that is, if a pupil has been performing well in previous assessment, it is assumed the pupil will also perform well in every assessment, mood of the examiner, extraneous factors such as handwriting legibility and ink color, adding up scores wrongly. Furthermore, in the event that the test is not assessed anonymously, the score may also be influenced by the grader's personal sentiments toward the student, the student's sex or race, or teacher expectations regarding the caliber of the student's work. Therefore, the child's future rests on the results of a single exam that is externally administered, graded, and has the potential to elevate or depress. The above mentioned factors are the issues on ground on the educational system associated with marking of theory-based examinations

Belo-Osagie (2014) stated that the main issue with manual essay and practical test scoring is its unreliability. The Senior School Certificate Examination (SSCE) examined the inter-rater reliability of oral English evaluation. The West African Examination Council (WAEC, 1993) discovered that ratings given by examiners to the same group of students varied significantly. This is one of the concerns we have with the issue of manual marking of theory-based questions. A student's performance on a theory examination is less likely to be an accurate indicator of their understanding of the content being evaluated since there is often disagreement among graders—or even among the same grader in the case of multiple readings—about the appropriate score for a particular theory question. This weakens the validity of the assigned score because it does not rightly reveal the level of subject mastery the student attained. Some of these factors calls for concern by stakeholders and as result of that, stakeholders such as WAEC, NECO and other examination bodies are looking for a

way to reducing or eliminating some of these challenges in manual marking either in internal examinations or public examinations, and as such, experts has suggested the use of automated system (Clark, 2014).

However, some educational institutions has started using automated system in multiple choice questions, but the one of theory-based questions is yet to be developed. In light of these constraints, the researcher decided it was essential to create an automated system that would grade theory-based test questions and ensure students mastery of the subject.

Theory-based examinations

Essay exams require students to show that they grasp the academic material covered in the course and that they can synthesis their mastery of it in a short amount of time (Ogunji, 2011; Amadi & Opuiyo, 2018). Essay exams are not the same as course assignments, which call for students to write well-researched essays with support from teachers, classmates, the annals, the digital materials, and other sources in due course. The composition essay, which is frequently accustomed to instruct and assess writing abilities, is not the same as the test essay (Ogunji, 2011; Amadi & Opuiyo, 2018). Unlike exams for courses that require essays, composition focuses on the writing process, requiring a lot of time and multiple drafts, as well as the development of complex cognitive skills (like inference, investigative, and decision making) related to communicating with an audience (Emaikwu, 2012). They have no indications or alternatives. The questions contain one or more subquestions. The two main ways they are commonly seen are in extended reaction and limited response. The first evaluates synthesis and problem-solving abilities. The answers are spontaneously generated and based on knowledge and understanding. For the second, there are certain requirements, such distinction and comparison.

Students must elaborate on their understanding and expertise in order to write their responses, which often include a few criteria. Exams are a tool used by educators, including instructors, to gauge how much knowledge pupils have retained throughout the school day. This can hold significance in many contexts based on the intended use of the outcomes. Determining the knowledge that pupils have acquired in the classroom is a crucial problem for the education sector (Amadi, 2018). Teachers want to be sure that they are utilizing an appropriate assessment technique that aligns with their learning goals when they evaluate what their pupils have learned. The essay examination, in contrast to best writing practices, is a first draft that has not been peer-reviewed, has not received feedback, and has not been permitted to be edited or proofread using outside resources. An first draft of the essay is produced as a result of the essay examination, which most likely do not fully or precisely depict the entirety of the student's write-up or even thought process (Onyibe, Uma & Ibina, 2015). However, the essay test still requires the student to provide well-reasoned written arguments that demonstrate their grasp of a subject matter. The essay examination's requirements, nonetheless, mean that giving written communication a lot of importance might lessen the validity of test results (Onyibe Uma & Ibina, 2015).

Because of the associated outcomes, essay assessments have begun to gain increasing traction. Questions in the essay style provide a higher quality of information than multiple-choice questions (Herandien, 2015). An essay test often consists of a few questions that the student must remember and arrange their information into coherent, cohesive responses to. There are two types of essay test items: extended response and short answer. They offer the chance to evaluate more complicated student traits and higher degrees of attribute attainment, claims Herandien (2015). An additional benefit of writing an essay is that, unlike multiple-

choice exams where students can guess the correct answers, the teacher can plainly see how much information the student has gained from what they have written. For a student who struggles with exams, writing an essay can be a far more effective method of assessing knowledge than using multiple-choice questions.

Essay examinations are not without their issues. Examining essays might be more difficult and less economical to administer. Multiple-choice examinations may now be graded using technology, and it takes a lot less time than essay assessments. Essays cannot be graded fast in boxes used for multiple choice examinations by a bubble sheet optical reader system. It is far more professional to utilize multiple-choice examinations rather than grade three hundred essays for a teacher with more than three hundred pupils. When required to prepare an essay for grading, students who struggle with writing may feel inferior. For a person with a learning handicap, this may be the case. Connolly & Cullen (2017) state that essay questions present a difficult task that requires written responses that might be anywhere in length from a few paragraphs to many pages. Like short response questions, they let students explain what they know and demonstrate creativity, but they also make it harder for them to make up an answer that passes muster. These questions may be made quickly and easily, but marking them can be difficult and take some time to reach grader consensus.

Despite its many flaws and restrictions, essay-style questions are still one of the most popular types of evaluation questions, according to Gronlund (1998). There are different types of essay questions: Unstructured-essay questions: These essay questions are structured to grant the writer as much flexibility as possible in defining the kind and extent of his response—subject, of course, to the intrinsic limitations given by the subject matter, the time allotted, etc. They

are sometimes referred to as free-response or extended-response inquiries.

Structured-essay questions: In contrast to unstructured essay questions, these questions—also referred to as restricted-response questions—give the test taker significantly less discretion over the type and extent of their response. In fact, they are purposefully written to provide the test taker with a great deal of direction on the format and subject matter of the essay.

According to Brereton (2014), while creating the question, students are given a set of tasks, that are well-defined and are aware of the particular expectations placed on them. Once more, this is crucial for both extremely explicit structured essay and short note questions as well as free answer questions. According to Nnam and Inah (2015), it is much harder to define the task exactly in the question's wording without negating the exercise's purpose by forcing the writer to follow a strict format and restricting their ability to choose, arrange, and present the information as they see fit.

One benefit of essay exams, according to Onyibe, Uma, and Ibina (2015), is that they allow students to show that they have achieved higher level skills like analysis and critical thinking. Students can exercise their own judgment, writing styles, and vocabulary through written assignments. Compared to other item types, they need the least amount of preparation time. Studies show that pupils study more effectively for exams that are essay-based than for selection (multiple choice) exams. Prioritizing broad topics, basic concepts, and interrelationships above precise specifics is how students prepare for essay assessments.

Essay assessments, according to Onyibe, Uma, and Ibina (2015), also allow the teacher a chance to remark on the students' development, the caliber of their thinking, the depth of their comprehension, and any challenges they may be facing. They give preference to kids who write well and are well-groomed. These present opportunities

for pupils who have a propensity to stray or misinterpret the question's basic idea. Essay exams provide students a challenge that is practical. Instead of answering multiple-choice questions, people in real life must organize and articulate their opinions.

Marking essay-type questions

According to Brereton (2014)'s book, one of the main drawbacks of essay exams is the high level of subjectivity that is unavoidably involved in the marking process, which can significantly lower the tests' dependability. For instance, the marker might find it challenging to eliminate from his assessment the following: his personal feelings; whether positive or negative; about the essay's author; his own personal opinions; or prejudices; about the subjects being discussed or the arguments or points of view being made; and extraneous details that have no bearing on the content itself such as messy presentation, hard-to-read writing, misspelled words, improper grammar, etc.

According to Brereton (2014), ensuring consistency in the scoring of essay assessments is crucial in order to award the same amount of points for equally good responses. Essay elements can be scored using one of two methods: the analytical or point technique and the holistic or rating method.

1. Analytic: Before scoring, prepare an ideal answer in which the major components are defined and assigned point values. Read and compare the student's answer with the model answer. If all the necessary elements are present, the student receives the maximum number of points. Partial credit is given based on the elements included in the answer. In order to arrive at the overall exam score, the instructor adds the points earned on the separate questions.

2. Holistic: This approach entails taking the student's response as a whole and evaluating its overall quality in relation to other students' replies or in accordance with particular standards you establish.

Essay exams are only reliable indicators of student success if students are prepared to take them. Many first-year college students lack the experience necessary to complete an essay test since they were not taught this ability in high school.

Automation in Education

Educators are already seeing some of the advantages of automation. Teachers no longer have to spend endless hours grading multiple-choice assignments by hand thanks to technological advancements that free them up to engage in more meaningful, one-on-one interactions with students instead of wasting time on such tedious tasks (Higgins & Heilman, 2014). Horbach & Zesch (2019) state that one important illustration of how automation and robots can allow educational institutions to accomplish more with less resources and budgetary restrictions is the time saved by something as basic as multiple-choice scanners. Nonetheless, several other domains within the educational system are beset by manual labor that would be more appropriate for automation. Automation has the potential to provide students with more autonomy and access by creating "self-service" environments.

Smart Theory Based Examination System

Smart Theory Based Examination System (STBES), is an automated system developed to mark theory examinations without any human interference. An automated system is made up of hardware and software that has been configured and built to function without the need for human input or guidance from an operator for each task (Sagar, 2020). With automated systems, you can keep an eye on your operations in real time, spot issues early on, and make swift corrections as needed.

The use of new technology in education, such as computer-based tests, to evaluate students' knowledge has generated a lot of excitement. But these assessments are usually based on objectives where students

are asked to choose from a list of options. The CBT exams do not really bring out the knowledge a student has acquired. They are not allowed to put down their own understanding in a well-crafted manner, which every educated person is expected to master.

Every system has goals and standards for performance. The purpose of the STBES was to lessen the prevalence of both lost examination scripts and incorrect exam scores following the test. Put simply, the automated method assists in giving student papers the proper grades. The creation of software services has been completely transformed by the application of artificial intelligence technologies. STBES will be very effective in scoring student exams quickly and efficiently. This system's automation for marking theory-based test questions speeds up the grading process for every paper, saving everyone's time as they wait for the results. The STBE technique eliminates the potential for manipulating results through extraneous deductions. Additionally, since the STBE system effectively handles paper verification, exam-conducting firms do not need to hire employees to do so (West 2016).

According to West (2016), the admin user is referred to as "Admin" in this study. Semantic models are used by the STBE system to produce similarity estimates for pairs of potential keyword combinations. The STBE system is able to do this by cross-referencing student responses with the keywords from the instructors' marking scheme. In addition, the instructor may add a number of additional synonyms to strengthen the keywords that stand for specific details that have to be included in an acceptable summary. The selection of keywords that include the background information for a specific answer is necessary for training the semantic models. This background information is necessary for the automated system to compare student responses accurately with the marking guidelines of the

instructors and to give the student user prompt feedback.

The answers that students provide are given semantic representations by the system, which then compares them to the teacher's or lecturer's marking guide standard. This system may be accessed by three primary

users. They are the administrator, the lecturer/teacher, and the student. The administrator's responsibilities include adding, editing, removing, assigning, and creating students as well as uploading subjects and courses as Shown in Fig 1.

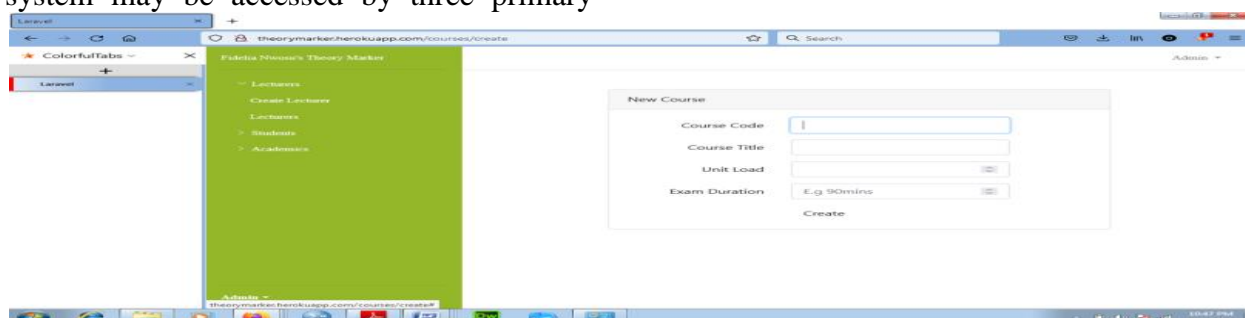


Fig 1: admin interface of the STBES

The task of teacher/lecturer is to upload, update and remove questions and also to check students result (as Shown in Fig 2).

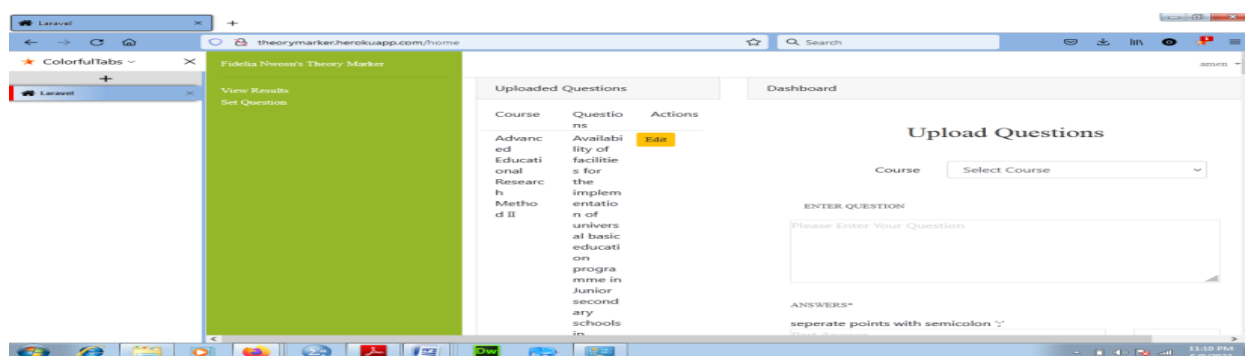


Fig 2: Lecturer interface of the STBES

The task of the student is to write exams and check results (as Shown in Fig 3).

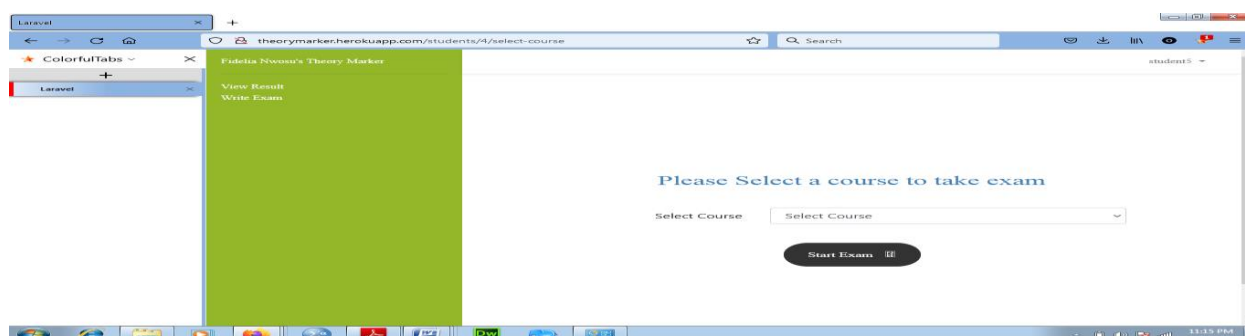


Fig 3: Student Interface of the STBES

The only goal of the user interface is to facilitate communication (interaction)

between the user and the machine. It is significant to note that only registered

users will be able to upload, amend, and remove questions as well as any conceivable keyword combinations. They will also be able to write and examine results by login into the system.

One essential component of the system is the back end. It is comparable to the system's brain, which is made up of the system database. An electronic collection of structured data, usually kept in an ordered manner within a computer system, is called a database. In order to construct this system, students will be able to react to questions posed by the examiners by putting their replies theoretically into the input box that is supplied. Once all of the questions have been answered, the answers will be marked and graded by an automated system. As the automated machine will mark theoretical questions, keywords submitted into the database will be accompanied with their synonyms. A term or phrase that is associated with a certain document or that summarises its contents is called a keyword. The instructor/lecturer, who is in charge of the questions and instructions, will upload these keywords.

Methodology

The methods used in conducting this research were presented under the following headings: study design, area/population of the study/instrument.

Study Design

Research and Development was the study's design. According to Gall, Gall, and Borg (2007), research findings are used in the R&D design process to create new products and procedures. These are then field tested, evaluated, and refined

using research methods until they satisfy predetermined standards for quality, effectiveness, or comparable criteria.

The Iterative Life Cycle Model, created by Royce (1970), is integrated into the creation of the automated system in this study. This research takes into account R&D design as its goal is to create an automated system for marking theory-based test questions.

Area/Population o/Instrument

The area of study is Nsukka LGA, Enugu state. The population for the study is 15 respondents which comprises of 15 Junior secondary III (JS 3) students. The selection of JSS 3 students was justified by the fact that they were unfamiliar with unified examinations and had no prior experience writing CBT exams. Exposing students to automated systems for theory-based test writing from the outset of unified exam writing is becoming popular (if accepted by the government). The instrument for data collection was the Smart Theory-based Examination System (STBES). The purpose of the automated system was for it to be tested for effectiveness, and any necessary upgrades may be made to the system using the information acquired. Before the students were introduced to the STBES to write precisely what was on the response script, they were given a mock examination on computer studies. This examination was written with pen and paper, and the instructor manually marked and scored it with the help of the marking guide that the teacher provided. The system score was gathered using this. The two scores were retrieved, and the manual scores and the STBES scores were compared.

Table 1: Showing Manual scores and STBES scores

Students ID	Manual Scores	STBES Scores
001	32	32.5
002	45.5	46.5
003	29.5	29.5
004	40.5	41.5
005	27	27
006	19.5	19.5

007	20	20
008	30.5	32.5
009	35.5	36.5
010	29.5	29.5
011	24.5	25.5
012	42.5	42.5
013	34	43.5
014	21.5	21.5
015	21.5	21.5

Table 4 shows scores of manually marked scores by teachers and Smart theory based examination system scores.

Findings

The finding shows that the STBES was effective and less time consuming and will help eliminates the possibility of guess work that may help the student pass such exams and lead to results that do not reflect the student's actual proficiency in the course. One of the main benefits of the STBES was its ease of use. Students praised it for making their exams so simple because, prior to submitting the exam, they could review the questions they had already answered and make changes without having to erase or cancel anything.

The power supply's epileptic characteristics caused the simulated exam to last longer than planned. The researcher offered to hire a generator that could power up to 20 computers at once in order to overcome this difficulty. The findings shows that, manual scores and STBES scores, were compared and it shows that STBES scores were found to predict Manual (teachers') scores.

Recommendations

The researcher hereby made the following recommendations:

- 1 In order to stop the practice of releasing theoretical test papers to untrained markers, Nigerian examination authorities should fully implement the usage of STBES for marking theory-based examination questions.
- 2 Examiners and other pertinent staff should have access to a variety of capabilities that might make automated marking easy and efficient.
- 3 Software developers could provide policies that will help in upgrading to a new STBES, and

relevant examination authorities might implement regulations that support the continued use of STBES marking.

Conclusions

This paper has developed a paperless Smart Theory-based examination system which allows the teacher/lecturer to assess the student subjectively, rather than suggesting answers to the student to choose from, as in objective computer-based tests. It provides for a more thorough assessment and eliminates the possibility of guess work that may help the student pass such exams and lead to results that do not reflect the student's actual proficiency in the course. The students do not need to fear based on any error that is associated to manual marking of scripts anymore. Apart from saving the school the cost of paper for exams, it also saves the space in offices, meant for the archival of past exam scripts. Soft copies of scripts are more secure and more durable, as they are backed up and kept in at least 3 separate locations. The cost of implementation is low, because it sits on existing computer network infrastructure. The only equipment needed, is the computer server which houses both the application and the database. Another good thing about this system is that students and teachers/lecturers using it will undergo remarkable improvement in computer proficiency. The result of the study demands the introduction of Smart theory-based examinations into the Nigerian Education Policy to help promote a more diverse and inclusive learning.

Reference

Amadi, E. C., & Opuiyo, A. R. (2018). Examination malpractice among Nigerian university students: A review. *International Journal of Innovative Legal and Political Studies*, 6(1), 13-17.

Belo-Osagie, K. (2014). *WAEC: E-marking of theory scripts begins*. The Nations. Retrieved from <https://thenationonline.net/waec-e-marking-of-theory-scripts-begins/>

Brereton, J. L. (2014). *The case for examinations. An account of their place in education with some proposals for their reform*. Cambridge, UK: Cambridge University Press.

Clark, N (2014). Impact of computer based test in Nigerian tertiary institutions: A theoretical view. *International Journal of Innovation Technology Integration in Education*, 1(1), 109 – 116.

Connolly, B., and Cullen. D (2017).“European court rules that exam scripts and comments constitute personal data.” William Fry, Dublin, December 21

Dong, F., & Zhang, Y. (2016). Automatic features for essay scoring—an empirical study. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing* (pp. 1072-1077).

Emaikwu, S. O. (2012). Assessing the impact of examination malpractices on the measurement of ability in Nigeria. *International Journal of Social Sciences & Education*, 2(4), 748- 757.

Gronlund, N. E. (1998). *Assessment of student achievement*. Allyn & Bacon Publishing, Longwood Division, 160 Gould Street, Needham Heights, MA 02194-2310; tele.

Harry, A, Raja, B. K, Kevin, M (2017),The implications of automation for education. Retrieved from <https://blogs.worldbank.org/education/implications-automation-education>

Herandien, P. (2015). *Educational software benefits for teachers and students in schools*. Cape Peninsula University of Technology. Retrieved from <http://www.researchgate.net/2023/02/dbms-note.html>

Higgins, D., & Heilman, M. (2014). Managing what we can measure: Quantifying the susceptibility of automated scoring systems to gaming behavior. *Educational Measurement: Issues and Practice*, 33(3), 36-46.

Nnam, M. U., & Inah, A. F. (2015). Empirical investigation into the causes, forms and consequences of examination malpractice in Nigerian institutions of higher learning. *International Journal of Novel Research in Humanity and Social Sciences*, 2(1): 52 – 62

Neil, N (2019). *Privacy's constitutional moment and the limits of data protection*. *BCL Rev.*, 61, 1687. Retrieved from <https://heinonline.org/HOL/LandingPage?handle=hein.journals/bclr61&div=44&id=&page>

Ogunji, J. A. (2011). Examination management and examination malpractice: The nexus. *Journal of International Education Research (JIER)*, 7(4), 53-64.

Onyibe, C. O., Uma, U. U., & Ibina, E. (2015). Examination malpractice in Nigeria: Causes and effects on national development. *Journal of Education and Practice*, 6(26), 12-17.

Sagar Shajahan (2020). Automation in the education sector – revolutionizing the world

of education. *Emperion Technologies* pp.
7-8

[Stephen J. Bigelow](#), (2022). *Cloud automation. What is cloud management? Everything you need to know. Cloud infrastructure design and management.*
Retrieved from
<https://www.techtarget.com/searchcloudcomputing/definition/cloud-automation?>

West, D. M, (2016). Big data for education: Data mining, data analytics, and web dashboards. *Governance studies at Brooking*, 4(1). 1-10.