ANALYSIS OF THE IMPACT OF THE STACK SYSTEM ON LEARNER ENGAGEMENT, PERFORMANCE, AND PERCEPTION ON ITS USE IN MATHEMATICS AT MASENO UNIVERSITY, KENYA

BY

JUMA ZEVICK OTIENO

THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE AWARD OF THE DEGREE FOR MASTER OF EDUCATION IN EDUCATIONAL TECHNOLOGY

SCHOOL OF EDUCATION

MASENO UNIVERSITY

DECLARATION

Declaration by the student

I hereby declare that this thesis is my original work and has not been submitted to any other institution of higher learning for a certificate, diploma, or degree academic credit.

Juma Zevick Otieno
MED/ED/00027/2019

SIGNED ……………………………………DATE ………………………………………

Declaration by the supervisors

This research thesis has been submitted for examination with our approval as university supervisors.

SUPERVISOR 1
Dr. Mildred. A. Ayere
Signature…………………………………. Date………………………………….
School Of Education,
Department of Communication, Technology & Curriculum Studies
Maseno University

SUPERVISOR 2
Dr. Michael.O. Oyeng'o
Signature…………………………………. Date……………………………….….
School Of Mathematics,
Department Of Pure and Applied Mathematics
Maseno University

ACKNOWLEDGMENT

I extend my gratitude to the Almighty for His unwavering guidance and support throughout my academic journey. This program has been a profound experience, filled with both enriching challenges and cherished memories. I would like to take this moment to express my heartfelt appreciation to my mother, Margaret Juma, and my brother, Russ Dan, for their unwavering support and encouragement during my studies.

I am also deeply thankful to my esteemed supervisors, Dr. Mildred Ayere and Dr. Michael Oyeng'o, for their invaluable guidance and unwavering support throughout my research journey. Their expertise and mentorship have been instrumental in shaping this project.

Furthermore, I extend my gratitude to the dedicated IDEMS team - David, Georg, Chiara, Esmee, and Santiago.

DEDICATION

To my mother, Margaret Juma, my source of strength and inspiration.

ABSTRACT

Feedback plays a crucial role in enhancing student learning outcomes. In large classes it is instrumental in maintaining student engagement, motivation, and overall academic performance by addressing misconceptions, and fostering deeper understanding of content amongst others. The School of Mathematics, Statistics, and Actuarial Science at Maseno University has had challenges with carrying out frequent formative assessments in its 20 high-enrollment introductory courses, due to large class sizes (500 to 1000 students) and limited lecturer support, affecting student engagement with content. In 2019, the faculty recognized the need for technological solutions to overcome the limitations of traditional assessment called the STACK system. The STACK system is a digital online platform with the ability to deploy grade randomized, advanced and adaptive mathematics related questions as well as provide immediate feedback, making it a valuable tool for large classes with diverse learning needs. However, despite its implementation, a research gap exists regarding its effectiveness as a solution for improving student engagement, performance, and perception in mathematics. The purpose of this study was to analyze the impact of the STACK system on learner performance, factors affecting engagement and learner perception on its use in mathematics at Maseno University. The research objectives were to: examine the correlation between learner test results within the STACK system and the scores in the end-of-semester exams, evaluate the factors affecting learner engagement with the STACK system and the end-of-semester exams and evaluate learner perceptions regarding the use of the STACK system as a formative assessment tool in mathematics. This study was guided by Gagne's theory of 9 levels of learning. This study was done at Maseno University, SMSAS. Population of the study was 4417 students enrolled in 5 high-enrolment courses already using STACK.Purposive sampling was used to select a single course for analysis which had 517 students enrolled. The study used STACK Quiz Analysis Guide, Students' Score Card, Online Survey Questionnaire, Interviews, and Focus Group Discussions (FGDs) to collect data. Cronbach's alpha ($\alpha = 0.87$) was used to check the survey questionnaire's reliability. Reliability and validity of the remaining tools were checked by experts in the field of education. Pearson's correlation analysis revealed a significant positive relationship between STACK scores and the end-of-semester exam ($r=0.63$, $p \leq 0.01$), as well as the frequency of STACK engagement and end-of-semester exam scores ($r=0.61$, $p \leq 0.01$). Key factors affecting learner engagement, are feedback, randomization, quiz availability, and peer interaction, having both positive and negative impacts on student learning and final outcome in the course. The online survey, interviews, and Focus Group Discussions (FGDs) revealed a generally positive perception of the use of STACK in formative assessment in mathematics. Despite this, concerns were expressed regarding insufficient feedback from certain STACK questions. These findings align with Gagne's theory of learning, emphasizing the significance of formative assessment, learner engagement, individual differences and positive perception, for effective learning. The study recommends further investigation to explore the impact of different types of feedback on learner cognition in undergraduate mathematics, to address the concerns raised by the students about insufficient feedback in certain STACK questions.

TABLE OF CONTENT

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND ACRONYMS

CAS – Computer Algebra System

CAT – Continuous Assessment Test

e-Assessment- Electronic Assessment

FGDs – Focus Group Discussions

SMSAS - School of Mathematics Statistics and Actuarial Science

STACK – Systems for Teaching and Assessment using Computer Algebra Kernel

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: INTRODUCTION

1.1 Background of study

Education technology, a transformative force in contemporary pedagogy, is reshaping the landscape of mathematics education (Alizadehjamal, 2022). Mathematics education is a critical aspect of academic development, fostering analytical thinking and problem-solving skills (Acosta-Gonzaga & Walet, 2018; Fatima, 2012; Maass, Geiger, Ariza, & Goos, 2019). However, instructing and assessing mathematics, particularly in large classrooms, poses significant challenges (Aina, 2022; Bethell, 2016; Cardoso, 2020; Kramer, Posner, Lawrence, Browman, et al., 2021; C. J. Sangwin, 2010). The conventional approaches to teaching and assessing mathematical skills often struggle to address the diverse needs of students and promote active engagement, thus affecting student perception and ability to acquire competent skills in the subject (Kramer, Posner, Lawrence, Browman, et al., 2021; Rowlett, 2011). In high enrollment classrooms, where the number of students presents logistical hurdles, mathematics educators grapple with providing effective formative assessments and timely feedback. Overtime, the integration of educational technology emerges as a promising solution, aiming to bridge the gap between traditional pedagogical methods and the evolving needs of contemporary learners (Alabdulaziz, 2021; Al-Hattami, 2020; Alizadehjamal, 2022; Alomran & Chia, 2018).

The landscape of mathematics education has undergone a transformative shift with the integration of digital tools in assessment practices such as Kahoot! which introduces a gamified dimension to assessments, turning quizzes into engaging competitions (Mdlalose, Ramaila, & Ramnarain, 2022), Wolfram Alpha, known for its computational power and its ability to provide step-by-step solutions (Abramovich, 2021), Maple, designed for advanced mathematics by tackling complex problems symbolically (Jones, 2008), amongst other tools. Systems for Teaching and Assessment which use Computer Algebra Kernel (STACK), is an education technology online tool integrated into the Moodle learner management system, which addresses the shortcomings of traditional assessment methods in mathematics and related subjects, by harnessing the power of computer algebra systems (C. J. Sangwin, 2002). It not only facilitates the generation of diverse mathematics questions but, crucially, provides instant feedback akin to the nuanced guidance a student might receive on paper. This distinctive feature sets STACK apart as a dynamic and rigorous tool, positioning it as a beacon in the realm of computer-aided assessment platforms compared to other platforms like Webster,Tutor Web, Quizziz, Kahoot, amongst others (Beliauskene & Yanuschik, 2021). The STACK system uses a distinctive grading algorithm called "potential response trees" to grade student answers provide feedback based on the mathematical properties of the student answer in comparison to the model answer. Figure 1.1 shows a basic STACK question with student response and feedback from STACK. This image was taken from a course using STACK at Maseno University.

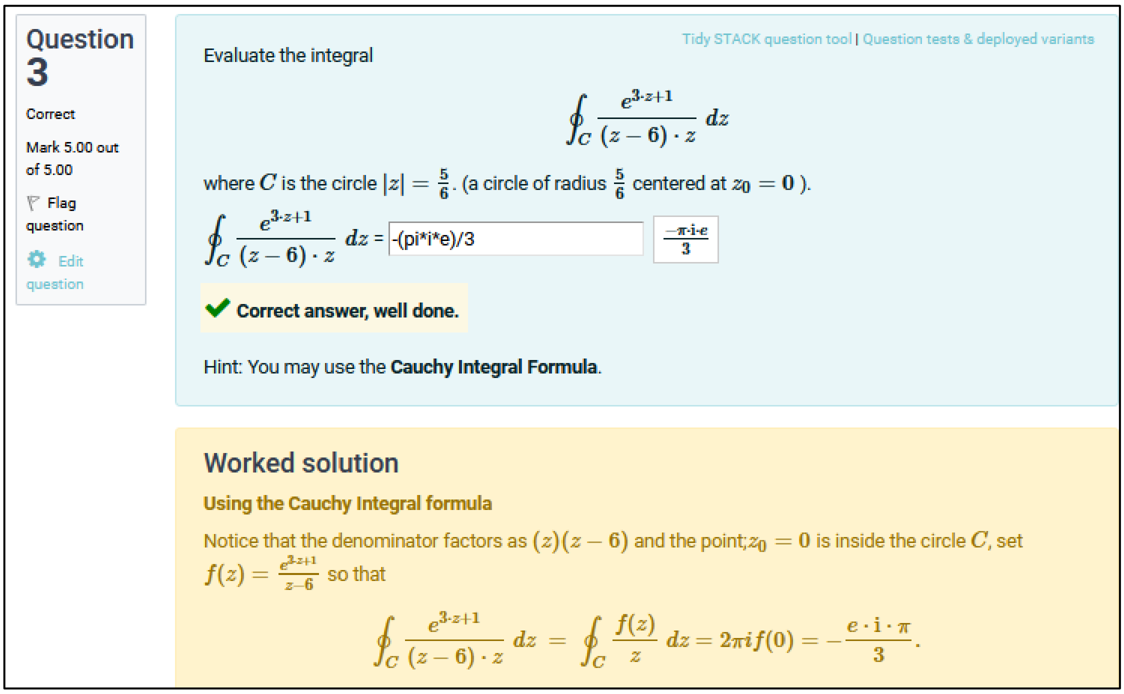


Figure 2.1. An example STACK question with feedback.

Figure 1.1 shows the STACK’s ability to discern and appreciate the varied approaches students may adopt in solving mathematical problems. Beyond a binary assessment of correctness, STACK's algorithm engages with the subtleties of mathematical reasoning, through targeted feedback on incorrect responses with misconceptions allowing for comprehensive and accurate assessment of simple to complex mathematical problems (Bach, Stephan, 2020; Derr, 2019; Knaut, Altieri, Bach, Strobl, & Dechant, 2022).

STACK's ability to provide personalized feedback and error recognition through its potential response trees hold the promise of improving students' retention and transfer of mathematical skills during practice with the randomized versions of the same problem in one sitting, resulting in enhanced performance in the subject (Barana, Marchisio, & Sacchet, 2021). Additionally, the system's capacity, in collaboration with the Moodle LMS, can track student engagement and performance which in turn assists in identifying areas where students may require additional support, enabling timely intervention to enhance learning outcomes something which is rather challenging in a traditional assessment setting with limited resources (Beliauskene & Yanuschik, 2021). According to Gagne and Gagne's (1985) nine-step model of learning, which encompasses gaining attention, informing learners of objectives, stimulating recall of prior knowledge, presenting content, offering learning guidance, eliciting performance, providing feedback, assessing performance, and enhancing retention and transfer, the importance of timely feedback in assessments cannot be overstated. Timely feedback enables learners to pinpoint gaps in their understanding before they escalate into substantial obstacles to learning (Gagné & Gagné, 1985).

According to Kallweit's (2019) cross-cultural analysis of various case studies on the use of STACK, there exists a statistically significant positive correlation between student performance in STACK and their final examination scores in comparison to paper-based assessments. Likewise, empirical studies have revealed that the implementation of STACK in formative assessment results is a remarkable enhancement in students' problem-solving capabilities and exam preparation. Notably, the correlation between STACK usage and students' final examination scores is significantly stronger in comparison to paper-based assessments (Knaut et al., 2022; Mäkelä et al., 2016; Tomilenko & Lazareva, 2020; Ustinova, Tomilenko, Imas, Beliauskene, & Yanuschik, 2020; Zerva, 2020).

While extant literature underscores the positive impact of computer-assisted assessment technologies, a critical examination of their efficacy in predicting learner performance remains limited to the European and Asian context. Notably, diverse scholarly works challenge the prevailing assumption that assessment technologies, particularly those reliant on computer-based methodologies like STACK, can adequately forecast performance in subjects demanding high-order thinking skills—a capacity traditionally associated with paper-based assessments (Kim, Belland, & Walker, 2018; Shute & Rahimi, 2017; Smolinsky, Marx, Olafsson, & Ma, 2020). The validation of the STACK system as a reliable and predictive assessment tool through the correlation analysis with final exam results is imperative for institutions contemplating its adoption or are already using it like Maseno University according to reports by Borio and Oyengo (2019). This validation would not only instill confidence in the reliability of the STACK system but also provides tangible evidence that investing in this technology aligns with enhancing overall subject mastery, thereby optimizing educational resources for improved learning outcomes. Furthermore, the study extends beyond this binary exploration by adopting a multifaceted approach. By delving into the broader landscape of technology adoption in classrooms, the research aims to investigate the reliability of STACK in varied learning contexts. Acknowledging the diverse challenges and opportunities inherent in different educational settings, as indicated in the works of Broekkamp and van Hout-Wolters (2007) and Chen, Xie, Zou, and Hwang (2020).

Engagement in education is a fundamental determinant of effective learning (Barana, Marchisio, & Rabellino, 2019; Trowler & Trowler, 2010). Engagement can be measured using either quantitative approaches such as learning analytics or qualitative approaches like interviews or focus group discussions (Henrie, Halverson, & Graham, 2015). For teachers, the ability to assess and ensure student engagement is paramount, as disengagement can hinder the learning process and impede academic progress. In mathematics, engagement is not just desirable; it is a necessity (Alizadehjamal, 2022; Bach, Stephan, 2020; Riske, Cullicott, Mirzaei, Jansen, & Middleton, 2021). Mathematics demands active cognitive involvement, regular practice with feedback to enhance problem-solving, and critical thinking. When students are actively engaged in mathematical tasks, they are more likely to grasp complex concepts, retain information, and transfer their learning to related tasks that demand the acquired cognitive skills (Deng, Benckendorff, & Gannaway, 2020).

Gagné and Gagné (1985), asserts that feedback plays a crucial role in promoting learner engagement. However, the challenge intensifies in settings with limited resources (Aina, 2022; Cardoso, 2020; Oyengo, Parsons, Stern, & Sangwin, 2021; C. J. Sangwin, 2002). In such contexts, technology becomes an invaluable ally. The integration of educational technology, exemplified by tools like STACK, becomes critical. Studies have demonstrated that in resource-constrained environments, technology can bridge gaps, providing interactive and dynamic learning experiences that enhance engagement according to Kallweit's (2019) report on cross-cultural analysis on the use of STACK across various contexts. Notably, the significance of STACK extends beyond mere technological integration. Drawing from Gagne's theory of learning, which emphasizes the importance of gaining attention, informing learners of objectives, and providing learning guidance, STACK emerges as a pivotal tool. Its capacity to offer personalized feedback, dynamic problem-solving environments, and varied question formats aligns seamlessly with Gagne's principles, fostering a learning atmosphere conducive to engagement.

Bach's (2020) posits a significant number of students having invested more than three hours per week engaging with STACK assignments, despite their full-time jobs and attending in-person lectures. Derr's (2019) study involving 2800 students showed improved exam scores due to engagement with STACK through formative assessments. STACK Mastery Quizzes, with randomized questions and unlimited attempts, enhance problem-solving skills in first-year mathematics courses according to Derr. Scholarly works such as Lowe and Mestel (2020), Nakamura, Taniguchi and Nakahara (2014), Oyengo, Parsons, Stern, and Sangwin (2015b ,2021), also demonstrate the effectiveness of STACK as a key intervention for enhancing learner engagement with feedback in mathematics, particularly in low resource settings.

While technology is being touted as a solution to promoting learner engagement, other schools of thought from various scholarly works such as Kramer et al., (2021), Riske, Cullicott, Mirzaei, Jansen, and Middleton (2021), and Sun, Guo, and Zhao (2020), disagree. These scholarly works argue that key drivers of student engagement with technology include motivation, the learning environment, the teacher's enthusiasm and involvement with the tool, as well as other factors. Understanding and studying learner engagement with STACK transcends technological considerations. It aligns with the broader goal of enhancing mathematics education, addressing resource limitations, and championing Gagne's principles of effective learning. Furthermore, the research gap that justifies the need to study factors that lead to various level/types/categories of learner engagement with technology and their impact on final performance lies in the limited understanding of the nuanced ways in which students interact with educational tools (J. Lee, Park, & Davis, 2018; Riske et al., 2021; Sun et al., 2020). Existing literature on STACK has insufficiently explored a comprehensive exploration of specific engagement categories and their distinct contributions to overall academic achievement. Investigating these nuanced aspects can provide valuable insights into the intricate relationship between engagement patterns and academic success, addressing a crucial gap in the current understanding of the dynamics between technology use, engagement, and learning outcomes. Therefore, the second objective of the study sought to uncover factors affecting learner engagement with the STACK system and its impact on their final performance.

Scholars across various disciplines all agree that students' perceptions of mathematics encompass a complex interplay of emotions, beliefs, and attitudes that significantly influence their engagement and performance in the subject (Chan & Wong, 2014; Chinofunga, Chigeza, & Taylor, 2023; Kayode & Anwana, 2023; Ogange, Agak, Okelo, & Kiprotich, 2018; Sikurajapathi, Henderson, & Gwynllyw, 2021). Generally, these perceptions are shaped by various factors, such as prior experiences, pedagogical approach, societal influences, and personal aptitudes. For many students, mathematics can evoke a range of emotions, from enthusiasm and confidence to anxiety and apprehension. According to Pokharel (2023), Yang (2013) and Noraini et al. (2018), students often view mathematics as a challenging and uninteresting subject, leading to a lack of motivation and interest. This perception is further reinforced by negative experiences such as poor grades and a lack of teacher support. Positive experiences, effective teaching strategies, and a supportive learning environment can contribute to favorable perceptions, fostering a sense of competence and enjoyment in mathematical pursuits. On the other hand, negative experiences, a lack of understanding, or a perception of mathematics as overly challenging can result in feelings of frustration and disinterest. Therefore, students' general perception of mathematics is a multifaceted phenomenon shaped by various factors. According to (Pokharel, 2023) recognizing and understanding these perceptions are crucial for educators to tailor their teaching approaches, create a positive learning environment, and address challenges that may hinder students' engagement with the subject.

According to various scholarly works that have delved into the multifaceted factors influencing the attitudes of students toward mathematics across all levels of education from primary to tertiary level, one key theme that recurs is the prevalence of poor attitudes, emphasizing the need to understand the diverse factors that contribute to negative perceptions (Gafoor & Kurukkan, 2015; Han & Liou-Mark, 2023; Holmes, 2015; Óturai, & Martiny, 2023). The introduction of technology, particularly tools like STACK, adds a new dimension to the discussion of learner perception in mathematics (Butcher 2008; Nakamura et al. 2012; Nakamura et al.,2014; Zerva 2020). Understanding how students generally perceive mathematics sets the stage for exploring how technology can be harnessed to positively impact these perceptions, creating a more inclusive and effective learning experience across all cultural contexts becomes imperative.

In Africa, persistent concerns about low levels of mathematics achievement prevail, as highlighted in the World Bank's flagship report, the World Development Report (WDR) 2018—Learning to Realize Education’s Promise (Carter-Rau & Olsen, 2018, 2019). The report emphasizes a learning crisis in global education, particularly affecting low-income and developing countries. In sub-Saharan Africa, less than 7 percent of students in late primary school are proficient in reading, while only 14 percent demonstrate proficiency in mathematics and the number keeps going down moving up the academic ladder. The challenges are further underscored by regional differences, with a 2014 assessment revealing that 58 percent of grade 6 students in West and Central Africa lack sufficient proficiency in reading. Moreover, the report identifies factors such as high enrollment, negative attitudes toward the subject, and the absence of well-designed student assessments as contributing to the persistently low levels of mathematics achievement in the region.

While empirical evidence indicates that the implementation of STACK may yield a beneficial impact on students' mathematical learner perception according to Derr, (2019), it is important to note the presence of a geographical disparity as there is a dearth of studies on the use of STACK in the African context. Secondly, a contextual gap is present as there is a need to examine the specific aspects of formative assessment with STACK that affect learner performance in mathematics at Maseno University, where the STACK system has been in use, since 2019 when it was first piloted in the African context, and how this informs learner perception of the use of STACK. Finally, a methodological lacuna becomes apparent as the majority of studies referenced in the literature review predominantly employ either qualitative or quantitative methodologies in isolation. This underscores the imperative for more robust research methods, specifically employing a mixed-methods approach, to comprehensively elucidate the influence of formative assessment using STACK on learner performance.

The challenges in mathematics education in high enrollment classrooms are not unique to Kenya; they resonate on a global scale (Aina, 2022; Bastedo, Altbach, & Gumport, 2023; Cardoso, 2020). Across the world, educational institutions are grappling with the implications of surges in student numbers for the quality of education, particularly in disciplines like mathematics. The demand for effective teaching methods and technologies has become more pronounced as institutions strive to maintain educational standards amidst increasing enrollments. Research on these technologies need to done as well for effective data-driven decision-making, flexibility, and resource optimization in the face of rising global competitiveness. The landscape of higher education in Kenya has undergone a transformative shift with a notable surge in student enrollment in universities. According to the Ministry of Education report (2021), there has been a substantial increase in total enrollment in public universities in Kenya. Specifically, the figures have risen from approximately 126,000 students in 2012 to surpassing 562,000 in 2020. This surge in enrollment, while indicative of increased access to higher education, has brought forth a set of challenges, particularly in mathematics education (Oyengo et al., 2021). Amidst this global challenge, this study narrows its focus to Maseno University in Kenya. The rationale for this specificity lies in the recognition that while the challenge is widespread, it manifests uniquely in different institutional contexts. By zooming in on Maseno University, the study aims to provide a nuanced understanding of how high enrollment impacts mathematics education in a specific Kenyan higher education setting. This targeted approach allows for a more in-depth analysis of the local dynamics, enabling the development of contextually relevant solutions that can potentially contribute to the broader discourse on mathematics education in high enrollment classrooms. Maseno University, therefore, serves as a microcosm through which we can glean insights that have broader implications for addressing this global challenge.

Table 2.1 Enrollment Statistics of Introduction Courses at SMSAS (2020-2022).

Course code	**Course Name**	**Enrollment (2020)**	**Enrollment (2021)**	**Enrollment (2022)**
MMA 100	Basic Mathematics	1350	1100	1000
MMA 103	Introduction to Linear Algebra I	879	900	700
MMA 101	Analytical Geometry	1200	1350	1200
MMA 102	Calculus 1	1400	1301	1378
MMA 200/215	Calculus II	1160	907	800
MMA 303	Complex Analysis I	900	700	517
MMA 404	Complex Analysis II	900	879	700

Note: *This table shows only a sample of courses with the highest enrollment recorded at SMSAS from the 20 high enrollment courses in the years 2020-2022 and is not an exhaustive list of all high enrollment courses. Only 5 high enrolment courses are using STACK for assessment, see section 3.4- Population of Study*.

Maseno University provides foundational courses for undergraduate students pursuing mathematics-related degree programs. These courses often accommodate a significant number of students per class, as illustrated in Table 1.1. The table presents a sample of the highest enrollment courses for the years 2020-2022, highlighting the large class sizes lecturers handle at Maseno.

IDEMS International (Innovations on Development, Education and the Mathematical Sciences), a non-profit Community Interest Organization plays a leading role in promoting development causes in education and mathematical sciences, with a particular focus on Africa (Oyengo et al., 2021). In 2019, IDEMS launched a project in collaboration with Maseno University to implement Computer Aided Assessment (STACK) in African universities. The selection of Maseno University as the focal point for this study is underpinned by its unique status as the initial pilot site for the introduction of the Systems for Teaching and Assessment using Computer Algebra Kernel (STACK) in the African context in 2019 according to Borio and Oyengo (2019). The subsequent study, however, distinguishes itself by placing a specific emphasis on learner engagement, employing a nuanced approach to measure and analyze behavioral, emotional, and cognitive engagement, extending data collection over a semester-long period for a more comprehensive analysis of the STACK system's sustained impact on learner outcomes, and incorporating a mixed methods research design to provide a more holistic understanding of the intricacies surrounding the integration of the STACK system in mathematics education at Maseno University.

1.2 Problem Statement

Technology integration as a solution to overcome the limitations of traditional pedagogy is a subject of diverse opinions, particularly considering the imperative to embrace it amid the escalating student enrollment in higher learning institutions like Maseno University. In introductory courses at Maseno, class sizes often soar between 500 to 1000 students per class in a single course, emphasizing the need for innovative teaching methods. The adoption of the STACK technology, piloted in five courses in 2019 with the support of IDEMS International, remains unexplored ever since, in the context of the 20 introductory mathematics courses at Maseno University despite its potential significance.

While proponents assert the STACK system's capacity to enhance teaching methods and elevate student outcomes, skeptics stress the demand for empirical evidence to establish its reliability in improving learner performance, engagement, and perception in mathematics, especially in low-resource settings. This ongoing debate concerning the integration of technology in education necessitates thorough research on STACK. Despite its promise, there is a notable absence of comprehensive investigations with empirical evidence in Maseno and the African context at large, where such insights are crucial for seamless upscale to be able to compete with the global community when it comes to mathematics education, inspite of the current limited resources within reach.

1.3 Purpose of the study

The purpose of this study was to analyze the impact of the STACK system on learner performance, factors affecting engagement and learner perception on its use in mathematics at Maseno University.

1.4 Objectives of the study

The specific objectives of the study were to:

(i) Examine the correlation between learner test results within the STACK system and the scores in the end-of-semester exams.

(ii) Evaluate the factors affecting learner engagement with the STACK system and the end-of-semester exams.

(iii) Evaluate learner perceptions regarding the use of the STACK system as a formative assessment tool in mathematics.

1.5 Research questions

The following questions were used to guide the researcher during the study;

(i) How does learner performance within the STACK system correlate with their scores in end-of-semester exams at Maseno University?

(ii) What are the factors affecting learner engagement with the STACK system at Maseno University?

(iii) How do learners at Maseno University perceive the use of the STACK system as a formative assessment tool in mathematics?

1.6 Significance

The significance of this study lies in its potential to provide insights into the use of STACK at Maseno University to stakeholders, such as educators (lecturers), policymakers, and students. With technology integration such as STACK being proposed as a solution. Amidst the ongoing discourse on the efficacy of technology in education globally, there exists a critical need for empirical evidence to substantiate its impact, especially in low-resource settings where there is little knowledge of its use.

Educators can benefit from the study's results by gaining insights into how to optimize the use of technology to provide regular formative assessments.

Policymakers can also benefit from the study's results by understanding how technology can be used to improve the quality of education in high-enrollment courses.

Finally, students can benefit from the study's results by experiencing improved engagement, performance, and change of perception in mathematics as pointed out by some of the literature in the background. The use of digital formative assessment tools like STACK can provide students with more opportunities to practice and receive feedback on their understanding of course material, ultimately leading to better academic outcomes.

1.7 Limitations of the study

The following limitations were considered when going through the findings to interpret the results correctly within the investigated context.

The study was conducted in only one university, Maseno University, which may limit the generalization of the findings to other universities in Kenya or similar contexts. To minimize this limitation, a large sample size of students from different departments taking mathematics courses at SMSAS was used, to increase the diversity of the study population. Additionally, a mixed-methods approach was used to provide a more comprehensive understanding of the impact of STACK, which can help to strengthen the generalizability of the study findings.

The study was constrained by limited resources, such as technical support and finances, which restricted the number of courses that could be included in the research. As a consequence, only one course from the population was selected for the study.

The study was conducted in a university setting where students had access to computers through the university library and departmental computer labs, as well as internet connectivity through the university WIFI, besides relying on their own devices like smartphones. However, the lack of access to personal devices outside the university setting and the limited technical know-how on the use of the STACK platform were still some of the limitations that the study encountered. To mitigate these limitations, the researcher, in collaboration with the course lecturer provided training sessions for the students on how to use the STACK platform effectively, including how to input answers and interpret the generated feedback from the system. This training was done using a dummy quiz with varied questions for them to practice on. Furthermore, when giving the STACK assessment to the students, the lecturer deployed the assessments such that students had a 7-day grace period to attempt the compulsory Test Quizzes which contributed to 50% of their continuous assessment score, while the Mastery Quizzes, which made up the other 50% of the score, were available throughout the semester.

1.8 Scope of the study

The study analyzed the impact of the Systems for Teaching and Assessment using Computer Algebra Kernel (STACK) on performance, learner engagement and perception in mathematics at SMSAS, Maseno University, Kenya. The study used a mixed methods research design, combining both quantitative and qualitative data collection methods.

The research focused on the MMA 303 Complex Analysis course at SMSAS, Maseno University, Kenya, with a total of 517 enrolled students. Data collection was conducted over a single semester of 10 weeks using multiple tools such as the STACK Quiz Analysis Guide, Students' Score Card, Online Survey Questionnaire, Interview guide, and Focus Group Discussion Guide.

Specifically, the study scrutinized the impact of STACK on key variables: such as performance in the mathematics assignments which were deployed using the STACK system (test results) and the final exam score. Learner engagement captured behavioral statistics such as attempts and scores within the STACK system and how these statistical measures reflected the students' outcome in the final exam score, while learner perception concerning the utilization of STACK in formative assessment was comprehensively examined through interviews, focus group discussions, and an online survey questionnaire. This approach enabled a nuanced exploration of the intricate dynamics surrounding the implementation of STACK in the selected mathematics course at Maseno University.

1.9 Assumptions of the study

The first assumption made in this study is that all the STACK quizzes used in the courses were appropriately designed for the learners in the course and aligned with the course content. This assumption is important because it ensures that the findings of the study are valid and reliable. If the quizzes used were too difficult or too easy for the learners, the results obtained from the study may not be accurate.

The methodological assumption of this study is that the use of both quantitative and qualitative methods, is appropriate for analyzing the relationship between learner performance in STACK and end-of-semester exam scores, evaluating student engagement and learning, and establishing learner perception of STACK as a formative assessment tool in mathematics respectively. The assumption is that the data collected through these methods are reliable and valid, and that the sample size is sufficient to provide meaningful insights.

1.10 Theoretical framework

Gagne and Gange (1985) proposed nine instructional events that are essential for effective teaching and learning in the classroom. These events include (1) gaining learners' attention, (2) informing learners of objectives, (3) stimulating recall of prior learning, (4) presenting stimuli, (5) providing learner guidance, (6) eliciting performance, (7) providing feedback, (8) assessing performance, and (9) enhancing retention and transfer of knowledge. The theoretical framework for this study was intricately woven into Gagne's Nine Events of Instruction, providing a structured and comprehensive approach to understanding the impact of STACK on mathematics education at Maseno University.

The introduction of the Systems for Teaching and Assessment using Computer Algebra Kernel (STACK) serves as the attention-grabbing element in this context. The integration of technology in mathematics education is designed to capture learners' interest and emphasize the significance of the upcoming learning experiences. The STACK system was used to deploy weekly assessments with the aim of assessing learning objectives for each and every concept taught in the course. This aligns with Gagne's emphasis on objectives measurement. The theoretical framework, in line with Gagne's model, recognizes the importance of connecting new learning experiences with prior knowledge. Learners engage with both the Mastery and Test Quizzes within the STACK system, recalling and applying mathematical concepts learned previously, fostering continuity in their learning journey. The deployment of STACK serves as the stimuli, introducing learners to interactive and technology-enhanced mathematics assignments. This step corresponds to Gagne's focus on presenting information to learners in a clear and engaging manner. Learner guidance is facilitated through the structured deployment of STACK quizzes. Learners receive guidance on using the system effectively, emphasizing the importance of clear instructions and support, as suggested by Gagne. Learner performance is actively solicited through their interaction with STACK. The system prompts learners to demonstrate their understanding of mathematical concepts, aligning with Gagne's step of eliciting observable responses. Immediate and constructive feedback is an integral part of the STACK system. This aligns with Gagne's principle of providing timely and informative feedback to learners, fostering a supportive learning environment. The assessment of learner performance, both within STACK assignments and traditional exams, corresponds to Gagne's step of formally evaluating the outcomes of learning experiences. The overall aim of the study, in line with Gagne's model, is to enhance the retention and transfer of mathematical knowledge. By exploring the impact of STACK on learner performance, engagement, and perception, the study seeks insights into how technology can contribute to sustained learning outcomes.

1.11 Operational definition of terms

Engagement - refers to the interactive participation of learners within the STACK system, it is quantified by the frequency of interactions within the STACK system, including the number of attempts made by learners, time spent on each quiz, and the overall performance in the quizzes.

Feedback- This is the numerical and textual comments displayed by STACK upon submission of students in each assignment.

Impact – refers to quantifiable indicators of engagement, such as the number of attempts, time spent, and correctness of responses. Furthermore, it also includes feedback and comments provided by learners within the STACK system, offering a more nuanced understanding of their experiences and challenges.

Mathematics- these are concepts that focus on numbers, space, patterns, and new conjectures to establish a truth by rigorous deduction from appropriately chosen axioms which are taught at SMSAS, Maseno University.

Perception- refers to propositions learners says concerning the use of STACK system in formative assessment.

Learner Performance – this refers to the test result scored by the student in either the STACK Quiz or the end of semester exams, which was a written exam.

The STACK system – which is also referred to as STACK/ STACK technology, is a digital tool with the ability to deploy high order thinking assessment, grade and give detailed feedback too.

CHAPTER TWO: LITERATURE REVIEW

2.1 Introduction

In this section, the literature was done thematically as per the objectives:

(i) Formative assessment with Technology and Student Performance.

(ii) Learner Engagement with Technology in Formative Assessment.

(iii)Student Perception on the use of Technology in Formative assessment.

2.2 Formative Assessment with Technology and Student Scores in the final exam

Formative assessment, a pedagogical approach focused on providing timely feedback to enhance learning, has witnessed a transformative shift with the integration of technology (Lane et al., 2019). Just to mention but a few, technology has substantially accelerated the feedback loop in formative assessment. This timeliness is crucial for students, providing them with the opportunity to address misconceptions or gaps in understanding promptly.

Technology has allowed for a diverse range of assessment modes beyond traditional paper-and-pencil methods. Interactive simulations, virtual laboratories, and multimedia presentations provide dynamic ways to gauge student understanding. This diversification aligns with the recognition that students have varying learning preferences and strengths. In addition, the integration of technology has enabled the creation of adaptive assessments that respond to individual learning needs. Intelligent algorithms can tailor questions based on a student's previous responses, ensuring that the assessment aligns with their current level of understanding. Educators can track progress, identify patterns, and make informed decisions on instructional adjustments. This data-driven approach allows for targeted interventions and the optimization of teaching strategies. Last but not least, interactive technologies engage students actively in the assessment process. Gamified quizzes, collaborative online activities, and interactive platforms turn assessment into a participatory experience rather than a passive task.

The marriage between formative assessment and technology presents a powerful synergy, promising dynamic opportunities to personalize learning experiences, engage students actively, and improve overall educational outcomes (Bhagat & Spector, 2017; D'Angelo, Rutstein, & Harris, 2016; M. J. Lowe & Vespestad, 1999).Various digital tools have emerged to augment formative assessment practices. Online quizzes, interactive simulations, and intelligent tutoring systems like Tutor Web and Kahoot provide educators with valuable insights into student progress, allowing for immediate intervention to address learning gaps (M. J. Lowe & Vespestad, 1999; Mdlalose et al., 2022). These tools promote active student engagement, self-assessment, and a continuous feedback loop. While the integration of technology in formative assessment brings substantial benefits, challenges persist. Issues related to accessibility, equity, and the need for teacher professional development warrant careful consideration. In addition, research on formative assessment with technology is imperative to evaluate the effectiveness of these tools, drive pedagogical innovation, ensure feedback quality, enable customization, address scalability, maintain global relevance, and contribute to continuous improvement in educational practices (B. Chen, 2015; de Haan, Vrancken, & Lukszo, 2011; Facer, 2011).

Contemporary scholarships have delved into the impact of various technology-enhanced formative assessment tools across different educational levels and subjects. The literature reveals a nuanced understanding of how these tools impact student learning, teacher practices, and the overall educational ecosystem. In the following sections, we will narrow the focus of the literature review to explore formative assessment with technology on student performance, engagement with content and perception in mathematics.

The current literature review underscores an expanding corpus of research that investigates the efficacy of formative assessment technologies such as Kahoot! which introduces a gamified dimension to assessments, turning quizzes into engaging competitions (Mdlalose et al., 2022), Wolfram Alpha, known for its computational power and its ability to provide step-by-step solutions (Abramovich, 2021), Maple, designed for advanced mathematics by tackling complex problems symbolically (Jones, 2008), and the STACK system, amongst other tools, in augmenting student performance within diverse educational contexts. Numerous scholarly works have delved into the impact of formative assessment strategies in the realm of mathematics education. Konert and Jansen (2018) conducted a systematic review of 44 studies and found that formative assessment had a positive impact on learner achievement in mathematics. Nieveen and Folmer (2019) also conducted a meta-analysis of 62 studies and found that formative assessment improved learner achievement in mathematics. However, both studies did not focus specifically on the use of the STACK system.

Kallweit (2019), study on the use of Moodle STACK in different European and Asian learning institutions found a stronger correlation between student performance in Moodle STACK and final exam results compared to paper assessments. Tomilenko and Lazareva (2020) support the idea that Moodle STACK enables learners to practice their problem-solving skills and enhance their learning by providing feedback, thereby preparing them effectively for exams. Similarly, Sangwin (2015) and Zerva (2020), have underscored a positive association between the utilization of formative assessment through STACK and students' final examination outcomes.. Mäkelä et al. (2016) investigated the use of STACK in providing assignments to students at Tampere University in Finland. Mäkelä et al. (2016) report that almost 90% of students attribute their improvement in the course to the use of STACK in formative assessment. Last but not least, Faber et al. (2017) investigated the effects of digital formative assessment on student achievement in primary school students and found a positive effect on student achievement.

Despite the growing body of research on the effectiveness of STACK, a knowledge gap and contextual gap exists in the context of Africa regarding the correlation between STACK performance and final exam scores in undergraduate mathematics courses. While the studies mentioned above provide valuable evidence of the effectiveness of STACK in enhancing student performance, they predominantly focus on European and Asian learning institutions. The context application of STACK is crucial, especially as an online platform, because it allows for the adaptation of mathematical content and assessment methods to specific educational contexts. Considering the diverse learning environments, resources, and challenges in different institutions, tailoring the use of STACK ensures its relevance, effectiveness, and meaningful impact on learner engagement and performance. This context-specific approach helps address unique educational needs, making STACK a versatile tool for enhancing mathematics education in various settings, which this study sought to uncover by first assessing its reliability when predicting learner performance in the final exams. Focusing on African universities is essential due to the specific educational context and challenges prevalent in the region (Bethell, 2016) . The unique characteristics of African universities, including high enrollment in mathematics courses and potential resource constraints, necessitate an investigation into the applicability and impact of the STACK system in this setting. Furthermore, the empirical gap in understanding the correlation between STACK performance and final exam scores in African universities underscores the need for region-specific research to contribute valuable insights, address existing gaps in the literature, and provide evidence that can inform educational practices and policies in the African context.

Notwithstanding the positive results on the impact of STACK mentioned by prior scholarly works, some scholars have presented a counterargument challenging the notion that assessment technologies such as Moodle STACK are more effective than traditional paper assessments in enhancing student learning. Beliauskene and Yanuschik (2021), de Haan, Vrancken, and Lukszo (2011), Reich (2021), Thomas (2022), and Ustinova, Tomilenko, Imas, Beliauskene, and Yanuschik (2020) argue that a student's academic performance is predominantly determined by intrinsic motivation to learn, the learning environment, and the psychological well-being of learners, thereby disputing the school of thought that assessment technologies like Moodle STACK are superior to traditional paper assessments in promoting student improvement.

Other scholars have also concurred with the aforementioned schools of thought (Fynn & Mashile, 2022; Guangul, Suhail, Khalit, & Khidhir, 2020; Jarrah, Alwaely, & Darawsheh, 2022; Peytcheva-Forsyth & Aleksieva, 2021). Furthermore, the reliability of electronic versus non-electronic evaluation methods in measuring student competency in STEM education has been a matter of significant contention. Several scholars, including (Guangul et al., 2020; Kayode & Anwana, 2023), have contributed to this discourse by emphasizing that the reliability of formative assessment technologies is questionable in STEM education. This is primarily because the nature of mathematics necessitates hands-on activities as opposed to merely tapping on devices and the complexities involved in precisely assessing students in these fields (Maass et al., 2019) . The scarcity of literature on the utilization of STACK in any other university within the Kenyan context underscores a significant gap in the current academic discourse. While discussions surrounding the efficacy of technology in education are prevalent as discussed in various scholarly works such as Alruwais, Wills, and Wald (2018), the specific application and impact of STACK in Kenyan universities remain unexplored. This research aims to fill this void by providing empirical insights into the adoption and learner performance when using the STACK system in a Kenyan university setting, thereby contributing valuable knowledge to the existing body of literature.

Investigating the correlation between STACK performance and final exam scores in undergraduate mathematics courses in the Kenyan context is important because it would provide empirical evidence on the reliability of the STACK system in enhancing student performance and assess whether it is a worthwhile investment that encourage student learning within an institution among other intrinsic factors.

2.3 Learner Engagement with Technology in Formative Assessment.

Although there may be some disagreement regarding its extent, educators and researchers are increasingly recognizing student engagement as a multifaceted construct that encompasses multiple aspects of learning activities (Lijie, Zongzhao, & Ying, 2020; Wang, Binning, Del Toro, Qin, & Zepeda, 2021; Watt & Goos, 2017; D. Yang, Lavonen, & Niemi, 2018). However, there is no consensus on the extent and dimensions of learner engagement in the context of educational technology. In this study learner engagement was used to refers to the active and interactive participation of learners with the STACK system, it is quantified by the frequency of interactions within the STACK system, including the number of attempts made by learners, time spent on each quiz, and the overall duration of engagement. This approach was to ensure that the concept is effectively captured and enables meaningful analysis in the context of the STACK system.

Behavioral engagement, which refers to the observable actions and activities of learners, has been stressed for its significance in fostering student participation and commitment to educational activities, according to Olivier, Galand, Hospel, and Dellisse (2020). Blumenfeld, Kempler and Krajcik (2006), and Corno and Mandinach (1983), highlight that cognitive engagement on the other hand, involves the mental effort and investment in understanding and learning, emphasizing its importance in students' intellectual involvement and active processing of educational content. Emotional engagement refers to learners' affective experiences, feelings, and attitudes or perceptions as mentioned in various scholarly writings, highlighting its connection to positive emotions and its role in influencing overall engagement levels (Codispoti, Mazzetti, & Bradley, 2009; Özhan & Kocadere, 2020; Sikurajapathi et al., 2021). The frequency of interaction with educational platforms and the time spent on tasks provide valuable insights into learners' commitment, focus, and the maintenance of a continuous and meaningful learning experience according works of (Codispoti et al., 2009; Derr, 2019).

Henrie, Halverson, and Graham (2015) highlight that engagement can be effectively assessed through either quantitative or qualitative methods, providing researchers with a range of tools to explore this complex phenomenon.

Quantitative methods often leverage learning analytics, utilizing data generated by educational technologies and learning management systems (LMS) (Bulger, Mayer, Almeroth, & Blau, 2008; Dickinson et al., 2021; Goodman et al., 2017; Ullah & Anwar, 2020). According to these researchers, tracking students' interactions with online materials, time spent on tasks, frequency of logins, and performance in assessments. Learning analytics offer a systematic and scalable way to gather large-scale data, allowing for statistical analysis and the identification of patterns that indicate engagement levels. This approach is particularly advantageous for researchers aiming to conduct extensive analyses across a broad student population. However, it is essential to note that quantitative measures might oversimplify the nuanced nature of engagement, providing an overview rather than an in-depth understanding. On the qualitative side, engagement can be explored through in-depth interviews and focus group discussions. These methods delve into students' experiences, motivations, and perceptions when interacting with education technologies (Abdool, Nirula, Bonato, Rajji, & Silver, 2017; Groccia, 2018; Halverson & Graham, 2019). According the aforementioned scholarly works, through open-ended questions, researchers can gather rich, context-specific insights into the various facets of engagement. Qualitative approaches allow for a more nuanced understanding of the subjective experiences of learners, providing depth and context to engagement patterns. However, it's crucial to acknowledge that qualitative methods are resource-intensive and might not be easily scalable for large-scale studies. The choice between quantitative and qualitative approaches to measure engagement depends on the research objectives, the specific dimensions of engagement under investigation, and the overall research design (Henrie et al., 2015). Often, a combination of both quantitative and qualitative methods is employed to achieve a more comprehensive and holistic understanding of learner engagement in educational settings (Jiao, 2015). This study used mixed methods, combining both quantitative and qualitative data to explore the nuanced of various factors affecting learner engagement with the STACK system.

There is growing evidence that student engagement is crucial for successful learning and teaching (Henrie et al., 2015; Trowler & Trowler, 2010). Higher levels of engagement have been linked to better learning outcomes (Dotterer & Lowe, 2011; Lee, 2014). Conversely, disengagement has been linked to lower academic achievement according to Nystrand and Gamoran (1991) and is linked to poor performance.

According to Gagne's nine levels of learning, enhancing retention after teaching is widely recognized as the most challenging level (Lijie et al., 2020; D. Yang et al., 2018). The aim of formative assessment is to track student engagement with the content and offer continuous feedback that both instructors and students can use to enhance their teaching and learning (Jiao, 2015). Formative assessments are particularly helpful in enabling students to identify their strengths and weaknesses and concentrate on areas where they need to improve. Furthermore, it assists students in developing various efficient learning strategies, as well as enhancing their self-assessment and peer-assessment abilities.

Biggs and Tang (2007), and, Schaeffer and Konetes (2010), mentions that teacher effectiveness is determined by their ability to encourage the students to use available learning resources to achieve learning objectives. This might not be possible in high enrollment environments with limited resources (Collaço, 2017; Gitonga, Gatere, & Mwaura, 2016; Jiao, 2015; Witkowski & Cornell, 2015). The main school of thought in the scholarly works of Collaço (2017), Jiao (2015), and Witkowski and Cornell (2015), is that there is passive learning in classrooms with high student enrollment. Al-Hattami (2020), Bahati et al. (2019), Holmes (2015),Nakamura et al. (2014),Ogange et al. (2018),Weigel et al. (2019), and Zainuddin et al., (2020) on the other hand, all encourage the use of technology in increasing learner engagement because it provides teachers with more tools to support students.

Hodgson and Pang (2012) investigated the learning experience of 104 students taking statistics in a degree program in Hong Kong, involving weekly online continuous assessments. The findings of the study by Hodgson and Pang (ibid), revealed that learners not only showed the regular commitment of time to tasks but also took active steps to find answers themselves with their peers. Kramer et al. (2021) used a mixed-methods design to investigate the impact of replacing non-electronic modes of assessment with electronic assessment on learner engagement with content in a quantitative survey. Kramer discovered a positive deviation of 0.31 in the number of times students spent learning Algebra. Mikes (2021) conducted a survey on 48 mathematics teachers to determine how formative assessment with STACK influenced learner engagement. The study by Mikes (2021) discovered that teachers preferred using simpler assessment approaches to maintain learner engagement which also makes their work easier to monitor. In this case, we'd say the STACK assessment is comparatively simplified because it compiles students' analytics in the course at every level in an easier-to-interpret grade book. Derr (2019) Lowe, and Hunt (2019), and Nakamura et al. (2014), conducted a number of case studies on mathematics students at universities in the UK and Japan, respectively, to investigate the use of STACK in mathematics. All the scholarly works point out that students are more engaged with STACK than with non-STACK assignments. Another gap identified in this section is a methodological gap, which relates to the lack of knowledge on how to measure and optimize learner engagement in mathematics when STACK is being used (Henrie, Halverson, & Graham, 2015; Kallweit, 2019; Maamin, Maat, & H. Iksan, 2022). Although the literature acknowledges the importance of learner engagement for successful teaching and learning, there is still a need for research to investigate the relationship between STACK and learner engagement in mathematics, and how this impact optimizes to enhance students' overall learning outcomes. This research gap requires further empirical investigation to develop effective strategies for measuring and optimizing learner engagement in mathematics.

In contrast to earlier assertions, other scholarly writings challenge the notion that technology operates as a sole determinant shaping students' engagement with educational content according to Lee et al. (2018), Riske et al. (2021), and Sun et al. (2020). Their findings suggest a more intricate interplay of factors, asserting that student motivation, the learning environment, and involvement are critical determinants influencing student participation within the learning space. This critical perspective underscores the importance of delving into the intricacies of discrepancies between students' engagement with the STACK system and their final outcomes in end-of-semester exams. By recognizing the multifaceted nature of engagement and its dependencies on motivational, environmental, and instructional factors, the study seeks to provide a nuanced understanding of how these factors intersect with technology use. In doing so, it aims to contribute valuable insights that extend beyond simplistic assumptions about the direct impact of technology on engagement, aligning with the evolving discourse in the field.

2.4 Student Perception on the use of Technology in Formative Assessment.

In this study, learner perception was used to refer to propositions learners says concerning the use of STACK system in formative assessment.

The investigation into learner attitudes toward technology has been significantly advanced by varioys extensive studies (Han & Liou-Mark, 2023; Jamil, 2012; VANNATTA, BEYERBACH, & WALSH, 2001). These scholarly writings emphasize the pivotal role played by perceived usefulness and ease of use in cultivating positive attitudes among learners. The perceived utility and simplicity of technology emerge as crucial factors contributing to a favorable disposition.

Moving to learner perceptions, Abdul and Kurukkan (2015), Ngware, Oketch, and Mutisya (2014), explored this domain, shedding light on the influential role of learners' subjective interpretations and evaluations of technology. These studies highlight that students' perceptions directly impact their willingness to integrate technology into their learning experiences, encompassing assessments of its advantages, disadvantages, and usability. To examining learner experiences with technology, (Dermo (2009) and Derr (2019), delved into the multifaceted nature of these encounters in educational contexts. Considering factors such as engagement, satisfaction, and impact on learning outcomes, these studies recognize the intricate connection between the effectiveness of technology in education and the quality of experiences students undergo while interacting with it.

While the general literature on technology in education lays the foundation, specific studies on STACK are relatively limited. Pioneering efforts by various scholarly works in the use of the STACK system in undergraduate mathematics, have delved into initial perceptions and experiences with STACK in specific contexts (Derr, 2019; Kallweit, n.d.; C. J. Sangwin, 2010; Zerva, Sangwin, Jones, & Quinn, 2022). These exploratory studies serve as precursors, unraveling the dynamics of incorporating STACK into educational practices and setting the stage for more in-depth investigations.

Formative assessment plays a pivotal role in gauging student progress and fostering effective learning experiences (Acee et al., 2017; Baya'a & Daher, 2009; Lim, 2019; Sikurajapathi et al., 2021). In recent years, there has been a notable shift towards technology-enhanced formative assessment, a transformation fueled by the rapid digitalization of education (Alabdulaziz, 2021; Al-Hattami, 2020; Facer, 2011; Johnson, Sondergeld, & Walton, 2019). However, within this evolving landscape, it is imperative to consider an often-overlooked aspect: student perception, especially in low resource settings where technology integration is considered the solution to the growing lack of resources for support.

The work of Vygotsky and Cole (1978) on the Zone of Proximal Development (ZPD) and socio-cultural theory provide a foundation for understanding how students construct mathematical knowledge and the role of technology in this process. In low-resource settings, where collaborative learning and scaffolding are essential, these theories gain added significance. Gagné's (1985) theory, known for its systematic approach to instructional design and its emphasis on the conditions of learning, was deemed particularly suitable for this study due to its focus on how individuals acquire and retain knowledge, when using STACK. In the context of understanding learner perception of STACK formative assessment in resource-constrained environments like Africa, Gagne's theory provides a structured framework to analyze the intricate interplay between technology, pedagogy, and learner experiences, which is vital for effectively addressing the unique challenges and opportunities presented by low-resource settings. The integration of technology into formative assessment has introduced both promises and complexities to mathematics education (Al-Hattami, 2020; Baleni, 2015; Beliauskene & Yanuschik, 2021; King, 2023; C. Sangwin, 2015b; Ustinova et al., 2020). While numerous studies have explored the impact of technology-driven formative assessment tools in mathematics, the majority of these investigations have concentrated on well-resourced environments for both in Africa and outside the continent, particularly studies relevant to the STACK community. This geographical gap becomes pronounced when considering the potential implications of STACK-enhanced formative assessment in within the African community and other similar context. By failing to address this gap, current literature leaves uncharted territory in the understanding of how technologies such as STACK impacts mathematics education in contexts with limited access to resources.

Existing research provides valuable insights into students' attitudes, preferences, and experiences in various educational environments where technology has been integrated (Bach, Stephan, 2020; Kallweit, n.d.; M. J. Lowe & Vespestad, 1999; T. W. Lowe & Mestel, 2020; Oyengo et al., 2021). However, there is a dearth of literature when it comes to understanding how students perceive the specific use of STACK technology in mathematics education, particularly in low-resource settings. This calls for an investigation into the unique factors, challenges, and benefits that characterize student perceptions of STACK in formative assessment within the context of undergraduate mathematics in resource-constrained environments.

Maseno University's pioneering efforts in integrating STACK technology on a large scale have not only set a benchmark for educational institutions across East Africa but have also opened up a wealth of opportunities and challenges according to the case study report by Borio and Oyengo (2019) and Oyeng'o et.,al (2021). This research, focusing on learner perceptions of STACK integration, not only has the potential to enrich the educational landscape at Maseno but also stands to influence pedagogical practices and policies throughout the East African region. By delving into the thoughts, attitudes, and experiences of students engaging with STACK, this study provides invaluable insights that can inform not only Maseno's ongoing efforts but also shape the broader discourse on technology-enhanced mathematics education in resource-constrained settings. As STACK technology continues to bridge gaps in mathematics education, this research contributes to its effective and sustainable implementation, ultimately empowering students and educators alike in their pursuit of mathematical excellence

CHAPTER THREE: METHODOLOGY

3.1 Introduction.

This chapter provides a detailed description of the research design, population of the study, sampling techniques used, data collection tools, their reliability and validity, the data collection procedure, data analysis approach, data presentation, and ethical considerations in the study.

3.2 Research Design.

This was a sequential mixed-methods study design which combined both correlational design and descriptive survey in it. The purpose of this approach was to complement the strengths of both designs, thereby providing a more comprehensive and holistic understanding of the research topic (Mujere, 2016). The first phase of data collection involved quantitative data collection and analysis, and then followed by qualitative data collection and analysis in that sequence.

3.3 Area of Study.

This study was conducted at Maseno University, which is situated in Maseno town, located along the Kisumu-Busia highway. The town is 25 kilometers from Kisumu, Kenya. The geographical coordinates of Maseno University are 0.0067° South and 34.5985° East.

The selection of Maseno University as the focal point for this study is underpinned by its unique status as the initial pilot site for the introduction of the STACK system in the African context (STACK Case studies, 2019). This historical precedence marks Maseno University as a trailblazer in integrating innovative educational technologies to address the challenges posed by high enrollment in mathematics courses. As the first institution to implement STACK within the African university landscape, Maseno University provides a distinctive and crucial case study that encapsulates the pioneering experiences, successes, and challenges associated with the initial adoption of this technology. By concentrating on Maseno, this study aims to offer a comprehensive and detailed analysis of the impact of STACK, thereby contributing valuable insights that can inform the broader adoption and implementation of similar technologies in other Kenyan universities and beyond. Lastly, the selection of Maseno was influenced by its active use of the Moodle systems in teaching and learning at undergraduate level (Ayere, 2022). STACK is already built within the Moodle system, therefore, this existing familiarity with both technologies, ensured a smooth integration into the university's educational ecosystem.

3.4 Population of the Study.

The targeted population of study comprised of 4417 students enrolled in 5 courses which were using STACK assessments in their course work. Table 3.1 provides a summary of the student enrollment for each of the courses as per the SMSAS 2022 data on student enrollment in high enrollment courses.

Table 3.4.1 Targeted Population of Study

Courses that are using STACK	**Class size**
MMA 100 (Basic Mathematics)	1000
MMA 103 (Introduction to Linear Algebra)	700
MMA 200/215 (Calculus II)	800
MMA 303 (Complex Analysis I)	517
MMA 404 (Complex Analysis II)	700
Total no. Students	4417

Note: Information retrieved from the School of Mathematics, Statistics, and Actuarial Science at Maseno University, 2021.

3.5 Technique of sampling and the sample size.

The study targeted a population of 4,417 students enrolled in five high-enrollment courses utilizing STACK assessments. Employing a purposive sampling technique, the research strategically selected the MMA 303 (Complex Analysis). Several practical considerations influenced this choice, including the accessibility of data, and logistical ease. The MMA 303 course, with an enrolment of 517 students, was deemed representative of advanced mathematical education at Maseno University (Palinkas et al., 2015). Furthermore, not all lecturers at Maseno University were uniformly committed to implementing the STACK system in their courses, leading to variations in data availability across courses. In light of these considerations, the study focused on the MMA 303 course to ensure a comprehensive and in-depth examination of learner performance, engagement, and perception within the context of STACK assessments.

3.6 Research Instruments.

Table 3.6.1 shows the instrumentations in this research, guided by the objectives of the study.

Table 3.6.1 Instrumentation as guided by the research objectives

Objective of study		Research Tool
i)	Examine the correlation between learner test results within the STACK system and the scores in the end-of-semester exams.	STACK Quiz Analysis Guide Student Score Card
ii)	Evaluate the factors affecting learner engagement with the STACK system and the end-of-semester exams.	Student Score Card STACK Quiz Analysis Guide Interview Guide Focus Group Discussion Guide
iii)	Evaluate learner perceptions regarding the use of the STACK system as a formative assessment tool in mathematics	Interview Guide Survey Questionnaire Focus Group Discussion Guide

3.6.1 STACK Quiz Analysis Guide.

Moodle tracks various statistics showing student behavior in quizzes. These statistics are; the number of times students attempted each quiz, their test scores on the weekly quizzes, and the length of time they attempted the quizzes (Gage, 2017). This study limited itself to the aforementioned statistics, which were compiled using the STACK Quiz analysis guide for correlational analysis. See Appendix 1 for more information.

3.6.2 Student Score Card.

According to Maseno University Examination Policy 2022, students must take the end-of-semester exam, which is a written exam that accounts for 70% of the total score in the course, according to university guidelines on assessment of undergraduates. The student score card served primarily as a student report card, with learner test scores on the exam and all weekly STACK Quizzes (calculated out of 30%). See appendix 2.

3.6.3 Survey Questionnaire.

The survey questionnaire consisting of ten items was used in this study to assess learner perceptions regarding the integration of STACK in formative assessment. Each item was evaluated on a five-point Likert scale, ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). Please refer to Appendix 4 for the specific survey items. It is worthy to mention that the questionnaire items were adapted from a previous study conducted by Jiao (2015), which explored learner perceptions concerning the integration of ICT in classroom assessment. It is important to acknowledge that explicit permission for the adaptation of these items was not sought. Nevertheless, the researcher made necessary modifications to tailor the questionnaire towards STACK, ensuring its relevance to the study context out of respect for intellectual property/work of the scholar. For example, one adaptation involved changing the original questionnaire item "I prefer online submission of Assignments" to "I prefer Continuous

Assessment Tests with STACK to non-electronic assessments i.e., paper assessment," thus aligning it more closely with the study's focus on STACK technology.

3.6.4 Student Interview Guide.

According to Hennink, Kaiser, and Marconi (2017), 16 to 24 interviews are needed to reach thematic saturation in a qualitative study. In this study, the researcher reached thematic saturation after interviewing 24 participants, randomly picked, to provide insights on the discrepancies identified in the STACK engagement data. The choice of interview questions was derived from the quantitative data in the online surveys, in the STACK quizzes, as well as the end of semester examination, check appendix 3. Engagement can also be measured using both quantitative approaches such as learning analytics as well as qualitative approaches like interviews or focus group discussions (Henrie et al., 2015).

3.6.5 Focus Group Discussion Guide.

Focus Group Discussions (FGDs) were the last stage of data collection. This came after the interviews, with the intention of corroborating the responses provided in the interviews, and allow students to challenge each other’s opinions on various topics of discussions, identified to having contributed to the quantitative data collected. See Appendix 5. The guide served as a structured outline that guided the conversations in the FGDs, ensuring that the researcher covered all the specific areas of interest and maintain consistency across all the 4 group discussions.

The size of the FGDs were guided by Gammie, Hamilton, and Gilchrist (2017) and Wilkinson (2011), who mentioned that ensuring the number of participants in the group do not exceed 12 is very important for quality data. Furthermore, diversification of the group members was done to ensure sufficient data collected. This study conducted 4 FGDs, with 9 participants in each group. The breakdown of the participants has been discussed in the introduction section in chapter 4.

3.7 Reliability.

Reliability is critical in educational research because it enhances the trustworthiness and dependability of the research tools (Mugenda & Mugenda, 1999).

To ensure the reliability of the study's tools, a pilot test of the online survey questionnaire was conducted with a separate group of individuals (50 students from another course, not part of the main sample). This pilot group represented 9.67% of the study's total sample of 517 students enrolled in the Complex Analysis 1 course. While the pilot group came from a different course, they were still part of the target population familiar with STACK technology. This allowed the study to assess if the questionnaire's questions were clear and understandable to potential respondents. As STACK was common to both courses, insights gained from the pilot group were transferrable to the main study.

Cronbach's alpha is particularly suitable for checking the internal consistency of a questionnaire because it checks the extent to which a set of items in the questionnaire measures the same underlying construct. This is essential when dealing with multiple items designed to capture multiple aspects of the same concept. In this study, the online survey questionnaire contained several questions that were intended to measure various aspects of the same construct related to learner perception on STACK adoption in undergraduate mathematics. Cronbach's alpha was well-suited to evaluate how consistently these questions were measuring that targeted construct. By calculating Cronbach's alpha coefficient, the study could determine if the questions in the survey were internally consistent and reliable in measuring the intended construct. The computed reliability coefficient of 0.87, as determined by applying equation 1, demonstrates a substantial degree of internal consistency within the questionnaire items.

The decision not to pilot the other data collection tools, such as the STACK Quiz Analysis Guide, Students' Score Card, Interview guide, and Focus Group Discussion Guide, was primarily due to practical constraints related to time and resources. Piloting these tools would have required significant additional time and effort to gather feedback and assess their reliability because they all required students to actively engage with the STACK system, analyze the performance/ grades and compare it to their final exam, then proceed with the interviews and FGDs. Reliability in this context was determined through alternative means such as a thorough review of the tools by subject matter experts or experienced researchers in the field. Additionally, efforts were made to ensure clarity and coherence in the design of the tools, aligning them with the study's objectives.

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum V_i}{V_t}\right)$$

Where;

α is the reliability coefficient (0.87)

k is the number of questions in the survey (10 items)

$\sum V_i$ is the sum of the variance of scores on each question (2.787)

V_t is the total variance of overall scores on the test (12.87)

Equation 3.7: Cronbach alpha reliability computation

3.8 Validity.

Both content and construct validity of the tools were strengthened through a thorough review by subject matter experts and experienced researchers in the field, ensuring alignment with the theoretical framework and research objectives. Mugenda and Mugenda (1999) emphasize the importance of assessing the validity of research instruments to ensure that each item accurately measures what it is intended to measure.

3.9 Data collection procedure.

Formative Assessment with STACK Assignments (1st Sept - 15th Nov 2022): The data collection process commenced with the administration of STACK assignments to students throughout a ten-week period as part of their regular coursework. These assignments served as formative assessment tools, aiding in monitoring student progress and offering feedback to enhance their learning outcomes.

Informed Consent (Throughout the Study): Prior to participating in the survey, students were presented with informed consent information. Participants were informed about the survey's objectives and the confidentiality measures involved in handling their responses, ensuring their voluntary participation. Electronic consent was obtained, and only students who provided consent were included in the survey.

End-of-Semester Online Survey (30th Nov 2022): After the completion of STACK assignments, an online survey was conducted at the end of the semester to collect data on students' perceptions of the effectiveness of formative assessment tools in enhancing their learning outcomes.

Final Exams (1st Dec 2022): Students sat for their final exams, which provided quantitative data on their performance in the course. Quantitative data was analyzed using descriptive statistics to identify patterns and trends in student performance.

Interviews (28th Nov - 30th Nov 2022): The second phase of data collection involved conducting 20 interviews with students enrolled in the course.

Focus Group Discussions (28th Nov - 30th Nov 2022): Finally, four focus group discussions (FGDs) were organized with students. Two of these FGDs were single-gender, while the other two were mixed-gender. The purpose of these FGDs was to gain more profound insights into students' experiences with the formative assessment tools and their perceptions of the tools' effectiveness in enhancing their learning outcomes.

3.10 Data analysis and presentation.

To examine the correlation between learner test results within the STACK system and the scores in the end-of-semester exams, Correlation Analysis was done on the Student Score Card and STACK Quiz Analysis Guide. Correlation statistics tables and scatter plots have been generated, offering a visual and quantitative exploration of the relationships between learner test results in the STACK system and the scores in the end-of-semester exams.

To evaluate the factors affecting learner engagement with the STACK system and the end-of-semester exams, thematic analysis, which involved transcription, coding, and grouping of responses from Qualitative data from the Interviews and Focus Group Discussions was done.

STACK Quiz Analysis Guide data was also subjected to Exploratory Data Analysis for further insights into various levels/cartegories of learner engagement with STACK. Extract presentations were created for responses from interviews and FGDs, providing insights into factors influencing learner engagement. Box plots and scatter plots visually represented the data gleaned from the STACK Quiz Analysis Guide.

The final objective, to evaluate learner perceptions regarding the use of the STACK system as a formative assessment tool in mathematics was measured using a survey questionnaire with 5 point liker scale and ten items as well as Interviews and Focus Group Discussions. Descriptive analysis/statistics were applied to the Survey Questionnaire findings, while Thematic Analysis was employed for qualitative data from Interviews and FGDs. A detailed table, showcasing statistics in percentages, summarized the student responses to the survey questionnaire findings. Extract presentations from interviews and FGDs were prepared to capture nuanced learner perceptions. The table 3.2 shows the summary of analysis and presentation used.

Table 3.10.1 Data Analysis and Presentation

Objective	Tool	Data Analysis method
(i) Examine the correlation between learner test results within the STACK system and the scores in the end-of-semester exams.	Student Score Card STACK Quiz Analysis Guide	**Correlation Analysis** **Data presentation**: Correlation statistics tables, Correlation scatter plots.
(ii) Evaluate the factors affecting learner engagement with the STACK system and the end-of-semester exams.	Qualitative data from the Interviews and Focus Group Discussions	**Thematic Analysis**. - transcription of recordings, coding of responses grouping similar responses. **Data presentation:** Extract presentation for responses from both interviews and FGDs.
	STACK Quiz Analysis Guide Student Score Card	**Exploratory data analysis** **Data presentation**: Box plots, scatter plot.
(iii) Evaluate learner perceptions regarding the use of the STACK system as a formative assessment tool in mathematics.	Survey Questionnaire findings	**Descriptive analysis/statistics** **Data presentation**: Table showing summary of statistics for each statement, in percentages.
	Qualitative data from the Interviews and Focus Group Discussions	**Thematic Analysis**. **Data presentation:** Extract presentation for responses from both interviews and FGDs.

3.11 Ethical considerations.

According to Haines (2017), Ketefian (2015) and Pearson et al., (2015), respect for individuals encompasses at least two ethical principles: firstly, the idea that individuals should be regarded as self-governing agents, and secondly, that individuals with reduced autonomy should be afforded safeguarding. Consequently, the principle of respecting individuals can be divided into two distinct ethical obligations: the obligation to recognize autonomy and the obligation to provide protection for those with compromised autonomy. In this study, the following ethical guidelines were followed:

1. Ethical clearance: Before proceeding to the field for data collection, the researcher applied for a research permit from Maseno University Ethics Review Committee (MUERC) through the School of Graduate Studies, see Appendix 9.
2. Informed consent: Before data collection, participants were informed about the research's purpose and their rights, including their right to refuse or withdraw from the study at any time. Participants were required to sign informed consent forms, indicating their willingness to participate in the study, see Appendix 6.
3. Confidentiality: Participants' identities were kept anonymous, and all data collected were treated with utmost confidentiality. The data were only accessible to the researchers involved in the study. The data collected in this study were protected against unauthorized access, alteration, or loss. All data were kept in password-protected computers, and only authorized personnel had access to them.
4. Anonymization of data: Anonymity of participants is another important ethical consideration that was observed in this study. The researchers ensured that the identity of the participants was kept confidential by assigning unique codes to the participants instead of using their real names. This was done to protect the participants from any potential harm or stigmatization that might arise from their participation in the study.
5. Voluntary participation: Participation in this study was voluntary, and no participant was coerced or forced to participate in the study.

CHAPTER FOUR: RESULTS AND DISCUSSION

4.1 Introduction.

This chapter presents the response rates proportion from the data collection tools and the findings of the study. The chapter begins with an overview of the response rates of the tools. It then provides a detailed presentation of the findings, including the statistical analysis used for the quantitative data and the method of analysis used for the qualitative data.

4.2 Participants.

Table 4.2.1 shows the summary of participant demographics.

Table 4.2.1 Participant Demographics

Faculty of origin	No. of students	Percentage out of the 514
Education	262	51.0%
Mathematics	216	42.0%
Business	26	5.0%
Physical and Biological Sciences	10	2.0 %
Total	514	100%

Note. Complex Analysis 1 (MMA 303) is a core course offered to students taking mathematics related programs from various schools within the university.

Table 4.2.2 summarizes the number of students who participated and the success level proportion calculated alongside.

Table 4.2.2 Participant Response Rates

	Expected participants	Observed participants	Success level Proportion (%)
Students who enrolled in the course and interacted with STACK the entire semester	517	517	100.0%
Students who sat the end of semester exams	517	514	99.4%
Online Survey Questionnaire responses.	517	350	67.7%
Interviews conducted	24	20	83.3%
Focus Group Discussions, with 9 participants each (4 Groups-1 male, 1 female, 2 mixed gender)	36	32	88.89%

Only 3 students did not sit for the end of semester examinations because they did not clear their school fees. They were excluded during the analysis. Of the 24 students expected, only 20 turned up for the interview, the remaining 4 could not make it due to unavoidable circumstances. Out of the 36 students who were selected, only 32 participants participated in the four Focus Group Discussions because 4 students pulled out when the FGDs began for personal reasons.

According to Mugenda and Mugenda (2003), a response rate of 60 percent or higher is generally considered sufficient for representative and accurate research results. However, it is vital to consider various study-specific factors like survey design and the characteristics of the target population that can affect response rates. In this study, rigorous sampling and survey design techniques were used to enhance participation and reduce non-response bias. As a result, the achieved response rate not only met but exceeded the recommended threshold, affirming the reliability and representativeness of the collected data

4.3 Correlation between Learner Performance in STACK and the End the final written exam.

The first objective of the study was to examine the correlation between learner performance in STACK and the end-of-semester exam score. In response to this objective, the STACK Quiz Analysis Guide, and the Student Score Card were used to collect data (see appendix I and II). Pearson's Correlation analysis was done to determine the relationship between; frequency of attempts on the STACK assignments (both mastery and test quiz), duration each student took to complete the weekly quizzes, the students' scores on the STACK quizzes, accounting for 30%, and their final exam scores, which constituted 70% of their overall performance.

Table 4.3.1 Correlation between STACK Performance and Final Exam Results.

		STACK_scores	Final_Exam
STACK_scores	Pearson Correlation	1	$.630^{**}$
	Sig. (2-tailed)		.000
	N	514	514
Final_Exam	Pearson Correlation **		1
	Sig. (2-tailed)	.630	
	N	514	514

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4.3.1 presents the relationship between learner performance in STACK and in the Final Exams. The table displays the correlation coefficients representing the strength and direction of the relationship between the two variables. The Pearson correlation coefficient quantifies this relationship. The correlation coefficient between STACK scores and Final Exam is 0.63 at .01 significance level, which is a moderately strong relationship. The positive value indicates a positive relationship, meaning that as STACK scores increase, Final Exam scores also tend to increase. The significance value (Sig.) represents the probability of obtaining the observed correlation coefficient by chance. In this case, the Sig. value is 0.000, which is less than 0.01. This indicates that with a 99 percent confidence level, the correlation between STACK scores and Final Exam is statistically significant. The sample size (N) for both variables is 514, indicating the number of participants or cases included in the analysis.

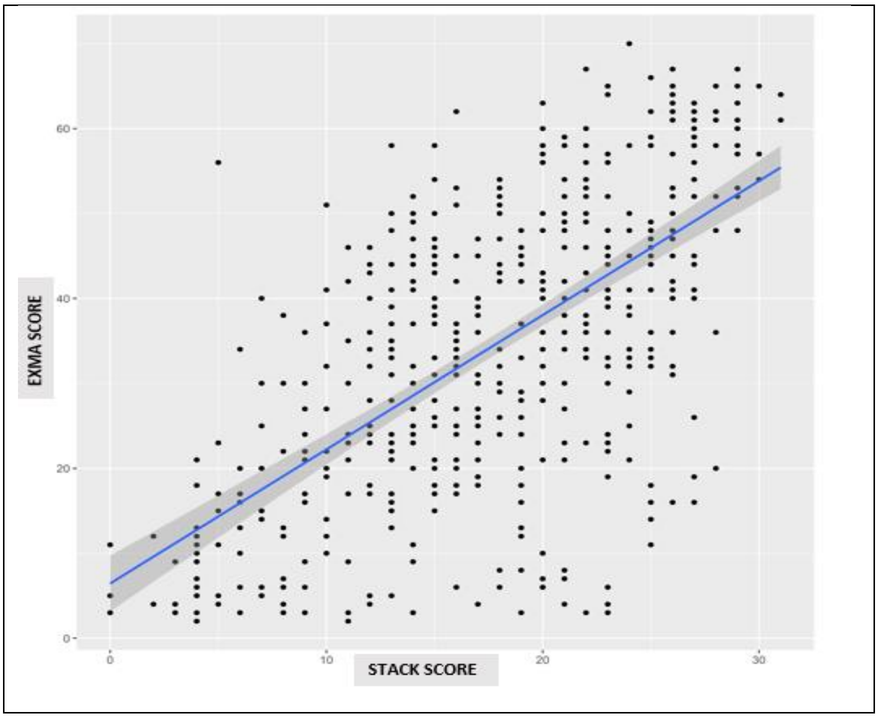


Figure 4.3.1 Student scores in the STACK Quizzes (out of 30%) and the final exam (out of 70%)

In order to further support and illustrate the findings from the correlation statistic table, a scatter plot of the data was generated, as shown in Figure 4.1. It can be observed that the data points are clustered closely around a straight line that slopes upward from left to right. The tight clustering of the data points suggests that as STACK scores increase, there is a consistent tendency for Final Exam scores to also increase. This provides additional evidence to support the statistical findings from the correlation analysis. Conversely, if the data points were scattered and spread out with no discernible pattern, it would indicate a weak or no relationship between the variables. However, in this case, the scatter plot confirms the presence of a clear positive relationship.

Table 4.3.2 shows the Pearson correlation statistics for the relationship between frequency of attempts on the STACK quizzes and final exam scores.

Table 4.3.2 Correlation between STACK Quiz Attempts and Final Exam Score

		STACK_attempts	Final_Exam_Score
STACK_attempts	Pearson Correlation	1	.612**
	Sig. (2-tailed)		.000
	N	514	514
Final_Exam_Score	Pearson Correlation **		1
	Sig. (2-tailed)	. 612	
	N	514	514

**. Correlation is significant at the 0.01 level (2-tailed).

This table presents the relationship between the frequency of attempts on the STACK quizzes and the scores in the Final Exams. The correlation coefficient between STACK attempts and Final Exam Score is 0.612 at .01 level of significance. This value indicates a statistically significant positive relationship, suggesting that as the frequency of attempts on the STACK quizzes increases, the scores in the Final Exams also tend to increase. The significance value (Sig.) represents the probability of obtaining the observed correlation coefficient by chance. In this case, the Sig. value is 0.000, which is less than 0.01 indicating the correlation between STACK attempts and Final Exam Score is statistically significant. The sample size (N) for both variables is 514, indicating the number of participants or cases included in this analysis.

Based on this table, the findings suggests that there is a statistically significant moderate correlation between the frequency of attempts on the STACK quizzes and the scores achieved in the Final Exams.

Figure 4.3.2, the scatterplot depicts the correlation between the frequency of attempts on the STACK quizzes and the corresponding scores achieved in the Final Exams. The scatterplot visually shows how as the frequency of attempts on the quizzes increases, there is a tendency for the Final Exam scores to also increase. This visual representation further supports the statistical findings presented in Table 4.3.2, highlighting the positive relationship between the two variables.

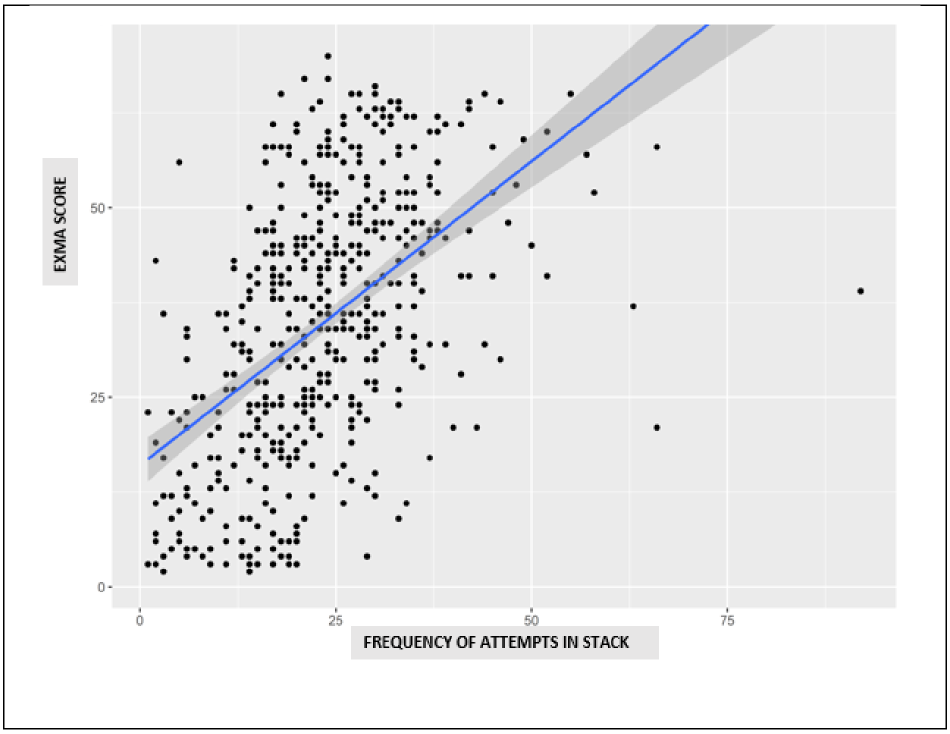


Figure 4.3.2 Frequency of attempts in STACK and the corresponding end of semester exam scores.

The study identified two limitations related to the measurement of the time students take with in the quizzes in Moodle. First, Moodle's duration recording does not capture the actual time students spend on quizzes because it only measures the time between opening and submitting a quiz. Second, many students worked on quizzes for multiple days because of the way the STACK Quizzes were implemented, making it challenging to interpret this data. These limitations made it difficult to use quiz duration data meaningfully in the study's analysis since there was no correlation of any kind between that variable and the final exam score in the course.

The quantitative analysis revealed a strong positive correlation between student performance in the STACK quizzes and their end-of-semester exam scores. This finding is consistent with previous research by Kallweit (2019) and Tomilenko and Lazareva (2020) and provides new empirical evidence for the potential of STACK in improving student performance in undergraduate mathematics education in Africa.

However, this research diverges from some earlier studies in the specific strength of this relationship. While the study found a moderately positive correlation, other researchers reported statistically stronger correlations (Knaut et al., 2022; Mäkelä et al., 2016; Tomilenko & Lazareva, 2020; Ustinova, Tomilenko, Imas, Beliauskene, & Yanuschik, 2020; Zerva, 2020). This discrepancy could be attributed to the following: Firstly, the level of significance set at 0.01 in our research was relatively narrow. Data points that would be deemed significant at a 0.05 level, for instance, could have been left out. This decision was made to minimize the risk of Type I errors in statistical inferencing (Lando & Mungan, 2018). However, this narrow level of significance may have led to the exclusion of some data points that could have contributed to a stronger correlation at .05 significance level.

Moreover, the presence of potential outliers in the data is another aspect to consider. Outliers' data points significantly different from the majority of the data, can have a substantial impact on correlation coefficients (Blllings & Voon, 1986; Burke, 1984). The decision not to eliminate outliers was due to their relevance in the analysis of learner engagement with STACK, which has been extensively discussed in the second objective. This approach acknowledges the potential value of outliers in providing insights into learner engagement and ensures that the analysis captures the full range of data points, including those that may deviate significantly from the majority (Modi & Oza, 2016). Lastly, the variations in the study population, including differences in prior exposure to technology-enhanced learning tools, could have introduced heterogeneity into the data (Beliauskene & Yanuschik, 2021). These variations are illuminated by the r-squared values ($r^2 = 0.36$), which was computed by finding the square of the correlation statistics (Senthilnathan, 2019) .

4.4 Analysis of Learner Engagement with the STACK system in Undergraduate Mathematics at Maseno University.

The second objective of this study was to evaluate the factors affecting learner engagement with the STACK system. To achieve this objective, multiple data sources were used, STACK Quiz Analysis Guide, 20 Interviews and 4 FGDs. Descriptive statistics was done on the STACK Quiz analysis guide to analyse learner engagement across the various quizzes. In this study learner engagement was used to refers to the active and interactive participation of learners with the STACK system, it is quantified by the frequency of interactions within the STACK system, including the number of attempts made by learners, the score achieved, time spent on each quiz, and the overall duration of engagement (Corno & Mandinach, 1983; Olivier et al., 2020).

The box plot in Figure 4.4.1 describes the variability and distribution of learner engagement, specifically focusing on the frequency of attempts made by students across various quizzes (Mastery and Test Quizzes, M1-M8 and T1-T8) within the STACK system. The size of the box plot reflects the extent of participation variability within each quiz, showing how students attempted and scored in terms of the grade achieved for each attempt. The box plots provide a visual representation of how data is distributed, offering insights into its central tendency, variability, and the existence of outliers. The vertical line inside the box represents the median, indicating the skewness of the data for each quiz. The larger the box plot, the greater the spread of the data, showing multiple attempts with varied scores achieved for each attempt. Conversely, a smaller box indicates less variability within the data. The study identified one challenge related to measurement of the time students take in the quizzes. Moodle could not measure the actual time spent by the students in each quiz, it only measures the duration when the quiz was opened and closed, therefore making it challenging to use that data meaningfully.

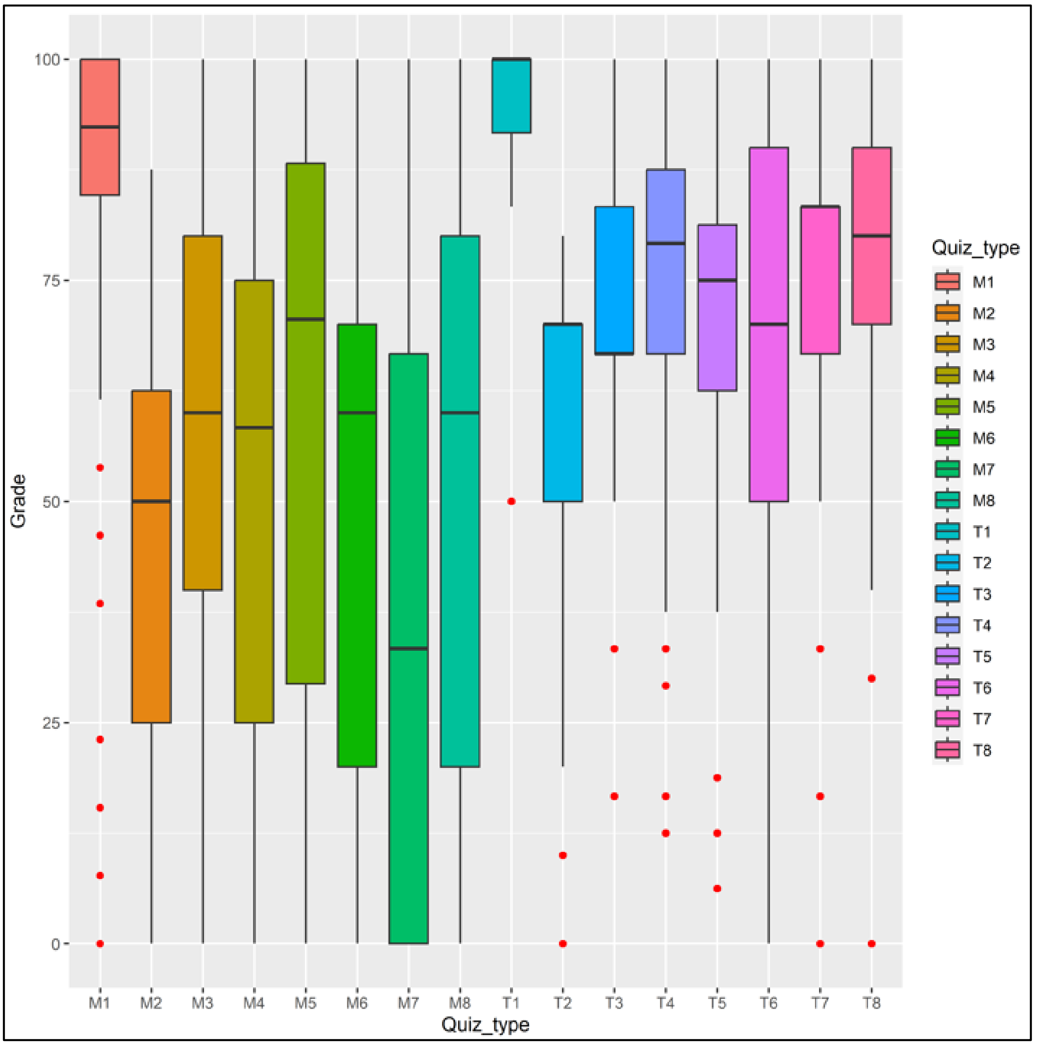
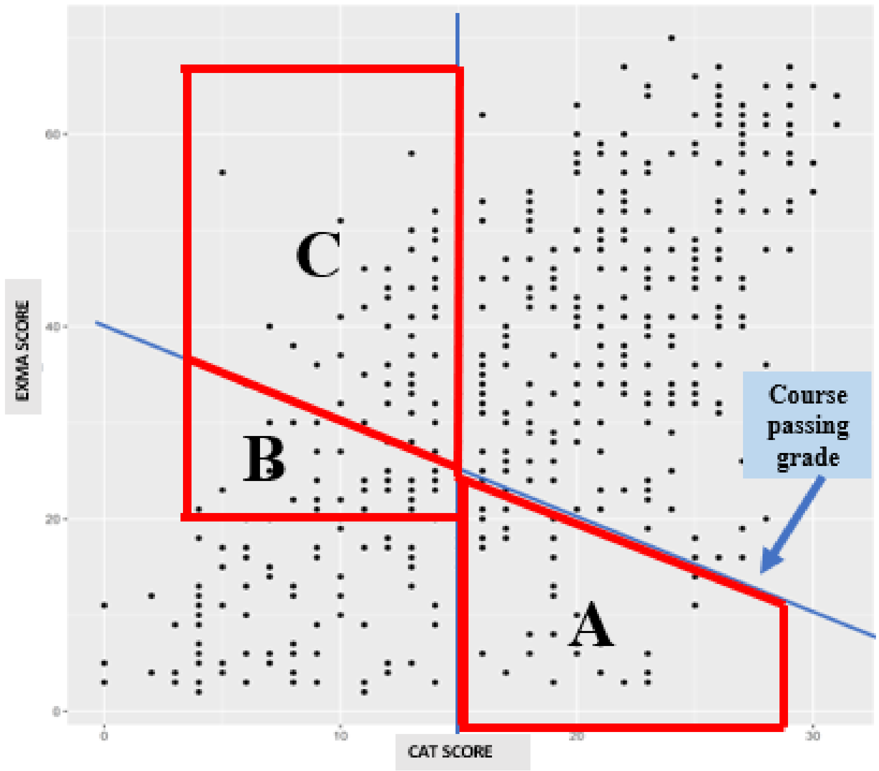


Figure 4.4.1 Boxplot showing Learner engagement in different Quiz types used in the course.

The quizzes were categorized into two main types: 8 Test Quizzes (T1-T8) and 8 Mastery Quizzes (M1-M8), both deployed on weekly basis. These quizzes assessed various content and were designed with a restrictive format, allowing students only a single attempt within a specified seven-day window. In contrast, the Mastery Quizzes were accessible throughout the entire semester, enabling multiple attempts by students. The box plots visually represent the distribution of learner engagement across different types of quizzes within the STACK system. Notably, the box plots for Mastery Quizzes (M1-M8) exhibit a wider and more dispersed spread compared to the corresponding Test Quizzes (T1-T8). This suggests that learner engagement in Mastery Quizzes is more varied, with a broader range of performance outcomes.

The wider spread in the Mastery Quiz box plots implies that students' engagement levels with these quizzes vary significantly. This could be indicative of diverse learning strategies, levels of understanding, or time commitments among students. The Mastery Quizzes, designed to be accessible throughout the semester and allowing multiple attempts, seem to encourage a more diverse range of engagement patterns. The study had a closer look at students who scored more than the average (15/30 marks) in the overall STACK system (CAT), and within this group, compared those who performed badly in the exam (≤15/70 marks) to the average. While the average CAT score of those students who performed poorly on the exam (20.6/30) was comparable to the average in the entire group (21.0/30), the study observed that most of the time, the students had only one recorded attempt per mastery quiz (median 10 mastery attempts in total, or 1.25 per quiz), while the average student in this group had multiple attempts (median 18 mastery attempts in total, or 2.25 per quiz).

To investigate the factors influencing learner engagement with the STACK system, exploratory data analysis was employed to classify students into three distinct categories based on their level of engagement, as illustrated in Figure 4.4.2. Category A (High Engagement, Unsuccessful Outcome). This group comprises students who demonstrated above-average engagement with the STACK quizzes, yet unfortunately failed the course (62 out of the 514). Despite their high engagement, these students faced challenges leading to an unsuccessful outcome. Category B (Moderate Engagement, Unsuccessful Outcome). Students in this category engaged with the STACK system but demonstrated below-average performance in the quizzes, ultimately resulting in course failure (51 out of the 514). Their moderate engagement levels did not translate into a successful outcome. Finally, Category C (Low Engagement, Successful Outcome). This category includes students who obtained below-average results in the STACK quizzes but successfully passed the course due to satisfactory performance in the final exam (257 out of 514).



Category A (High Engagement, Unsuccessful Outcome). This group comprises of students who demonstrated above-average engagement with the STACK quizzes, yet unfortunately failed the course

Category B (Moderate Engagement, Unsuccessful Outcome). Students in this category engaged with the STACK system but demonstrated below-average performance in the quizzes, ultimately resulting in course failure.

Category C (Low Engagement, Successful Outcome). This category includes students who obtained below-average results in the STACK quizzes but successfully passed the course due to satisfactory performance in the final exam.

Figure 4.4 2. Categorization of level of engagement in the course using STACK and Exam data.

Note. The diagonal line in the figure corresponds to the passing grade for the course, which was set at 40% in accordance with the university's policy for evaluating undergraduate students.

Investigations into factors influencing learner engagement were derived from a combination of qualitative data obtained through interviews and focus group discussions (FGDs). The thematic analysis of responses gathered from 20 interviewees and participants in the four FGDs revealed that feedback and randomization emerged as key themes influencing varied learner engagement with the STACK system. Feedback and randomisation came out as one of the main factors contributing to varied learner engagement with the STACK system. 15 out of 20 interviewees and most participants in the four FGDs agreed that STACK provided immediate feedback and randomized questions.

One interviewee remarked, *"... I really like how Moodle STACK provides instant feedback on different randomized versions of questions I did. Even when I get a question wrong, I could immediately see where I made a mistake and how I can improve… it helped me a lot in preparing for exams… although there were some instances when the feedback assumed some computational steps and it would take me a moment to actually understand what was done on the calculation...".* In spite of a majority of STACK questions having good feedback, there were instances when the feedback had technical breakdown within the coding in some key question which was raised as an issue as to why some quizzes had diverse attempts with learners trying their level best to make sense of the limited information, they can get their eyes on. One response from the FGD pointed this out elaboratively,

"Well, it is true that some questions were not easy and it is because each time we attempted it the feedback missed some steps in the explanation. For instance, when you look at the step-by-step calculation, you get lost halfway and so many students were getting discouraged in through quizzes while a few others kept repeating those quizzes with the hope of getting it right through guess work... ukiona kuna quizzes watu walifail sana – Translation- When you check, you'll find that people failed in some quizzes due to this.*"*

The second theme that came out from the qualitative analysis was availability of both types of quizzes (Mastery and Test Quiz) with a lot of emphasis on the role of the Mastery Quiz in supporting extensive learner engagement, through out the semester even when the test quizzes were closed. One interviewee stated, "*Although I could not do the test quiz, I found the mastery quizzes to be really helpful in preparing for the final exam. The fact that they were available throughout the semester meant that I could practice and improve my understanding of the course material over time. Also, the feedback provided after each question was useful in highlighting areas where I needed improvement.*"

This was backed up by the FGD responses. One FGD participant said, "*I know a few friends who were unable to complete some of the test quizzes due to unforeseen circumstances, but they made sure to utilize the mastery quizzes to prepare for the final exam. They were able to practice and understand the concepts better with the instant feedback provided by the system. When the final exam came around, they were more confident and performed well, despite their initial difficulties with the test quizzes."*

Peer interaction and collaboration came out as another strong contributor as to the diverse learner engagement; however, this came out as an opportunity that was negatively taken advantage of by some students, particularly those that dint pass the course despite having done so on the Continous Assessment test, which was a sum of all the STACK Quizzes computed out of 30 marks. One interviewee candidly shared, "*You know, doing the STACK Quizzes can be quite challenging. There are moments when you need to seek assistance from someone who's smart in the subject to complete the assignments for you...*" Another interviewee added, "*I won't name names, but I'm aware of some classmates who resorted to this (cheating). It is a risky move, but they still opted for it.*" These remarks portray the likelihood that a significant number of students devised strategies to exploit the opportunity and cheat instead of using the STACK system to practice with the quizzes and learn from it.

To investigate this phenomenon, the study focused on students who scored above average on the STACK quizzes. Those who performed poorly in the final exam (scoring 15 or fewer out of 70 marks) were compared to the overall class average. The investigation revealed that students who scored poorly in the final exam and engaged in academic dishonesty typically had only one recorded attempt for each mastery quiz, with a median of ten attempts in total or 1.25 attempts on average per quiz across all the 10 mastery quizzes. On the other hand, most students in this group tried the quizzes several times, usually around 18 times in total or about 2.25 times for each quiz. This finding supports the "*cheating theme*" drawn from the interviews. It is worth noting that not all students who didn't do well in the final exam cheated. Factors like exam-related anxiety and inadequate preparation, as mentioned by some participants, could also have contributed to poor exam performance.

Approximately 20% of the responses, both in interviews and FGDs, raised concerns about certain questions lacking detailed feedback on the specific computational steps leading to the final answer. This lack of clarity made it challenging for them to grasp the content, resulting in repeated attempts on the quizzes. It's important to note that this criticism is not directed at the STACK platform itself but rather at the formulation of certain questions. As one FGD participant expressed, "*The feedback from the system wasn't consistently helpful; at times, it merely restated the question without offering a clear explanation of the underlying concept.*" Nevertheless, this critique regarding the system's clarity in specific questions and computational steps is a valuable aspect to consider for platform enhancement. Additionally, three interviewees disagreed with the majority's perspective regarding the correlation between excelling in the STACK quizzes and failing the course. They emphasized that factors such as exam anxiety and inadequate preparation could also contribute to poor performance in the final exam.

Access to device and internet also came out as a contributor as to why some students did not engage extensively with the STACK system to score their optimum grades they should have. However this was dismissed by some students in the interview and a majority of students from the FGDs who mentioned that motivation was also key when engaging with the STACK system, inspite of the challenges at hand. One interviewee stated, *"...some students simply completed the tests to achieve the minimum score required in each Quiz to just to get it over with and move on… progress...*" This statement suggests that some students may have lacked intrinsic motivation to engage with the quizzes and improve their learning outcomes.

One of the issues raised by participants in this study is the issue of technological accessibility. Some students encountered challenges in fully engaging with the STACK system. This aligns with findings from previous studies that highlight disparities in access to technology, which can exacerbate educational inequalities (Corno & Mandinach, 1983; Derr, 2019; Henrie et al., 2015; Kramer, Posner, Lawrence, Krier, et al., 2021; Olivier et al., 2020; Özhan & Kocadere, 2020; Riske et al., 2021; Sun et al., 2020). While smartphones are prevalent, not all students have access to them or the necessary data bundles (Warschauer & Matuchniak, 2010). This can result in unequal opportunities for engagement with technology-enhanced learning tools. In addition, the study's findings identify specific areas for improvement in the design and implementation of STACK, particularly with regards to cheating and the lack of adequate feedback in some STACL questions which inturn affected how learners engaged. These findings inform future developments of the STACK system, as well as other similar e-learning platforms.

4.5 Learner Perception on the use of STACK as a Formative Assessment Tool in Mathematics

The final objective of the study was to learner perceptions regarding the use of the STACK system as a formative assessment tool in mathematics. The assessment of students' perceptions was done using an online survey questionnaire, the 22 interviews, and 4 FGDs. This section starts by presenting the quantitative findings from the survey then back them up with qualitative analysis from the interviews and FGDs. The online survey was administered as a google form. The Likert scale used in this survey consisted of 5 points ranging from Strongly Agree to Strongly Disagree, with a rating of 5 for Strongly Agree and 1 for Strongly Disagree. The proportion of student responses in each statement was calculated by taking the number of students who selected that statement and dividing it by the total number of respondents before multiplying it by 100%.

Table 4.5.1 Summary of the positive statements on learner perception on STACK Assessment

Positive Statements on STACK Assessment	**SD (%) 1.0**	**D (%) 2.0**	**N (%) 3.0**	**A (%) 4.0**	**SA (%) 5.0**	**Mean $\bar{x}$**
1. I prefer Continuous Assessments with STACK to non-electronic assessments i.e., paper assessment	12.9% (45)	9.7% (34)	9.1% (32)	14.9% (52)	53.4% (187)	3.86
2. There were Adequate self-assessment tests from the STACK quizzes which were enough to prepare me for the Exam	1.1% (4)	0.6% (2)	1.1% (4)	36.6% (128)	60.6% (212)	4.55
3. Having engaged with STACK through practice with feedback I now feel confident with solving problems in the related content	2.6% (9)	5.1% (18)	5.4% (19)	27.4% (96)	59.4% (208)	4.36
4. There was enough guidance in the feedback to help me understand the steps outlined in the STACK Quizzes	2.6% (9)	5.4% (19)	5.4% (19)	40.9% (143)	45.7% (160)	4.22
5. Using STACK in assessment makes the learning of mathematics more enjoyable and meaningful	5.4% (19)	0.0% (0)	8.0% (28)	27.7% (97)	58.9% (206)	4.35
Total Mean						4.27

Table 4.5.2 Summary of the negative statements on learner perception on STACK Assessment

Negative Statements on STACK Assessment	**SD (%) 1.0**	**D (%) 2.0**	**N (%) 3.0**	**A (%) 4.0**	**SA (%) 5.0**	**Mean x̄**
6. Having used STACK for Continuous Assessment, I still would not recommend its adoption over paper assessment in mathematics	40.9% (143)	26.9% (94)	13.7% (48)	5.4% (19)	13.1% (46)	2.23
7. I prefer paper assessment to STACK assessment in mathematics at Maseno	48.6% (170)	21.1% (74)	11.4% (40)	13.4% (47)	5.4% (19)	2.06
8. STACK doesn't provide adequate self-assessment tests to students while preparing for the end of semester exams	52.0% (182)	29.7% (104)	2.6% (9)	10.6% (37)	5.1% (18)	1.87
9. There were instances when feedback was not helping me at all	0.0% (0)	0.0% (0)	2.6% (9)	38.3% (134)	59.1% (207)	4.57
10. Given an opportunity, I would still not prefer Continuous assignments through STACK	40.3% (141)	28.3% (99)	2.9% (10)	7.1% (25)	21.4% (75)	2.41
Total Mean						2.63

KEY: Strongly Agree (SA)=1; Agree (A)=2; Neutral (N)=3; Disagree (D)=4; Strongly Disagree (SD)=5

The survey results indicate that a majority of students (68.3%) preferred continuous assessments through STACK over paper assessments, with 53.4% strongly agreeing with this preference. During the interviews, one student mentioned the following, "*I prefer continuous assessment with STACK because it allows me to identify my mistakes early and correct them before the final exam*.", indicating that the student has a positive preference to it. The 4 Focus Group Discussions revealed a consistent within the student body with one participant mentioning this *"I personally believe that STACK should be integrated. Before, when we didn't have it, many of us would procrastinate until the Continuous Assessment Test (CAT) was right around the corner, and then we'd cram everything just to pass the exam. We have realized that this approach doesn't work* *well with mathematics. Math demands consistent practice with feedback, and STACK compels us to do that on a weekly basis, which is why we prefer it to paper assessment."*

According to the survey questionnaire analysis, a significant proportion of students (60.6%) strongly agreed that self-assessment tests from STACK quizzes were adequate for exam preparation, while only a small proportion (5.1%) indicated that they were not enough for exam preparation. Student voices echoed this sentiment, as one interviewee emphasized, *"STACK forces us to practice regularly, which is important in mathematics. It is better than cramming everything just for the sake of passing the exam."*

Practicing with feedback from STACK improved students' confidence in solving problems, with 59.4% of respondents strongly agreeing with this statement. However, 97.4% found instances when feedback from the STACK quizzes was less effective indicating a need for improvement in this area. Qualitative insights from both the interviews and 4 FGDs underscored the positive impact of STACK feedback, with one student from the interview expressing, *"I like how STACK gives immediate feedback, so I know if I'm doing something right or wrong. It is like having a tutor with me all the time."* Concerning the accuracy of STACK feedback, which was a key concern that was aired out during the discussions one participant pointed out, *"I think STACK needs to improve on the feedback it provides. Sometimes it is not very helpful because it doesn't take into account the different ways we approach a problem."* This statement highlights the concern students had and the need to ensure formative assessments done through the STACK system is constantly reviewed.

Only a minority of students (13.1%) recommended retaining paper assessments in mathematics, while a majority of respondents (40.9%) strongly disagreeing with this statement. One response from the interviews highlighted that students who preferred paper assessment to STACK, did so because of lack of access to devices and being a lot of work. One student in the interview mentioned this, *"STACK is good, but I have concerns about implementing it right now. Not all students have smartphones, and this could pose a problem when it comes to completing weekly assignments if we were to adopt it."* The other response from the FGDs was as follows,

“STACK keeps students engaged all the time, which is not a bad thing. We do invest a significant amount of time in completing STACK quizzes. Now, imagine if all our courses incorporated STACK, that would significantly increase our workload. I guess (Speaking in swahili) hatukuja hapa kufanya hesabu peke yake, tuko na courses zingine pia tunafanya.

English translation - After all, we didn't come here just to study mathematics; we have other courses to attend to as well.”

Furthermore, only 5.4% of the students preferred paper assessment over STACK assessment, with nearly half of the respondents (48.6%) strongly disagreeing with this sentiment. Regarding the guidance provided by feedback, the majority of respondents (45.7%) found it adequate to help them understand the computational steps in the STACK quizzes. However, a small proportion of students (2.6%) found feedback unhelpful in this regard. Most respondents (58.9%) reported enjoying and finding meaning in learning mathematics with STACK assessments. One interviewee mentioned this, *“I like how STACK gives immediate feedback, so I know if I'm doing something right or wrong. It is like having a tutor with me all the time...”*

There is a mixture of how learners perceive the use of technology in formative assessment (Baya’a & Daher, 2009; Ogange et al., 2018; Walker, Topping, & Rodrigues, 2008; X. Yang, 2013). However, this study revealed nuanced insights into students' perceptions regarding the use of STACK in formative assessment in mathematics at Maseno University. While there are positive aspects highlighted, on learner perception with the technology as mentioned in other scholarly works (Beliauskene & Yanuschik, 2021; Sikurajapathi et al., 2021; Walker et al., 2008), negative feedback indicated areas of concern. Notably, a proportion of students expressed reluctance to recommend the adoption of STACK over traditional paper assessments, however a small proportion. Their responses are valid in any case one is considering to integrate STACK system in their teaching and assessment, this should be taken into account, not all students will prefer it. Additionally, dissatisfaction with specific aspects of the feedback system and an overall preference for paper assessment were highlighted by a significant percentage of students (97.4%). This shows the challenges when using technology in formative assessment as it requires consistent review for improvement (Beliauskene & Yanuschik, 2021; Rowlett, 2011).

These findings provide valuable empirical evidence that goes beyond a binary evaluation of the tool, offering a deeper understanding of students' perspectives. The identified concerns, especially regarding feedback and preferences, present opportunities for improvement in the design and implementation of STACK for continuous assessment in mathematics. This information is crucial for shaping future research endeavors and refining practices to enhance the effectiveness and acceptance of technology-based formative assessment tools like STACK.

The overall analysis which can be drawn from the findings in this section aligns with the body of knowledge that there is a general positive perception on the use of STACK in formative assessment in mathematics at Maseno ; Stephan, 2020; Knaut et al., 2022; T. W. Lowe & Mestel, 2020; Mäkelä et al., 2016; Nakamura et al., 2012; Nakamura, Taniguchi, & Nakahara, 2014).

CHAPTER FIVE: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

Chapter 5 of this thesis offers a summary of the study's findings, conclusions of these findings based on the researcher's perspective, and recommendations together with suggestions for future study. The study aimed to explore three main issues: (1) the correlation between learner performance in STACK and the end-of-semester exam score, (2) factor affecting learner engagement with the STACK system and (3) learners' perceptions of using STACK as a formative assessment tool in mathematics.

5.2 Summary of findings

This study supports Gagne's theory of instructional design, which highlights the importance of formative assessment, engagement, and feedback in improving student learning, perception, and performance in mathematics. The study found that implementing formative assessment through the STACK had a positive impact on student engagement and performance in mathematics, confirming Gagne's principles in a practical setting. The following subsections provide a summary of the outcomes of the research, organized by objective and corresponding conclusions.

5.2.1 Correlation between Learner Performance in the STACK system and the End of semester exam scores.

The findings of the study reveal a statistically significant, positive correlation between learner performance in STACK quizzes and their final exam scores, with the test results in the STACK system correlating at 0.63 and frequency of attempts within the STACK system correlating at 0.612 with final exam scores, both at a 0.01 significance level. The research differs from some earlier studies in the specific strength of the correlation, possibly influenced by the stringent significance criterion, the inclusion of outliers for insights into engagement, and variations in the study population's exposure to technology-enhanced learning tools. These nuanced findings contribute empirical evidence to the reliability of the STACK system in enhancing undergraduate mathematics education in low resourced settings, emphasizing the need for a comprehensive understanding of engagement factors and the impact of technology on diverse learner populations.

5.2.2 Analysis of Learner Engagement with the STACK system in Undergraduate Mathematics at Maseno University.

The study found that the way quizzes, especially the Mastery Quizzes available all semester, were designed led to different levels of engagement and outcomes among students. When we sorted students into different engagement categories, it showed diverse patterns, highlighting the need for a careful understanding of how students interact with the system. Key factors affecting engagement included feedback, randomization, quiz availability, and peer interaction. These factors had both positive and negative impacts on how students approached their learning. Instances of cheating and the need for clearer feedback in some questions were identified as areas that could be improved. The study also acknowledged challenges in students' access to technology but emphasized that motivation played a crucial role in engagement. In conclusion, these findings provide useful insights for improving the design of technology-based learning tools like STACK, addressing issues related to fairness, feedback clarity, and factors that motivate students.

5.3.3 Learner Perception on the use of the STACK system as a Formative Assessment Tool in Mathematics.

The quantitative analysis, based on survey responses, revealed a majority preference (68.3%) for continuous assessments through STACK over traditional paper assessments. Interviews and focus group discussions provided qualitative insights, with students expressing positive sentiments about the tool. Some highlighted the benefits of identifying mistakes early and the regular practice facilitated by STACK. However, concerns were raised, with a significant portion (97.4%) indicating instances where feedback was not effective in key questions that appeared in the final exams. Some students were reluctant to recommend STACK over paper assessments, citing challenges like lack of device access and increased workload. Despite these reservations, the overall perception of STACK in formative assessment was positive, aligning with broader literature on technology-enhanced learning tools. The study emphasizes the need for continuous improvement in feedback mechanisms and considers diverse student preferences in the integration of such tools in mathematics education.

5.3 Conclusions

Based on the study's results, the following conclusions can be drawn for each of the three objectives.

5.3.1 Correlation between learner performance in the STACK system and the end of the semester exam.

The study concludes that there exists a statistically significant, positive correlation between learner performance in STACK quizzes and final exam scores.

5.3.2 Analysis of Learner Engagement with the STACK system in Undergraduate Mathematics at Maseno University.

Key factors affecting engagement, are feedback, randomization, quiz availability, and peer interaction, having both positive and negative impacts on student learning, according qualitative data from interview responses and focus group discussions. Engagement with STACK was crucial for students when they needed a platform for practice with feedback. Areas for improvement, such as clearer feedback in certain questions and addressing challenges in students' access to technology, are identified. This study underscores the crucial role of motivation in student engagement.

5.3.3 Learner Perception on the use of the STACK system as a Formative Assessment Tool in Mathematics.

A majority prefer continuous assessments through STACK over traditional paper assessments. However, concerns about the effectiveness of feedback in certain questions and reluctance to recommend STACK over paper assessments highlight areas for improvement. The overall positive perception aligns with broader literature on technology-enhanced learning tools, emphasizing the need for integration in situations where traditional pedagogy poses some limitations.

5.4 Recommendations

The study proposes the following recommendations.

5.4.1 Recommendations to the School of Mathematics on the Use of STACK as per the objectives of the study.

(i) Given the observed statistically significant, positive correlation between learner performance in STACK quizzes and final exam scores, the study recommends the School of Mathematics to integrate STACK quizzes more systematically into the curriculum. Educators should emphasize the importance of consistent engagement with STACK as it correlates positively with exam performance and can be used to predict student final exam score to a greater extent. Continuous monitoring and analysis of student performance in STACK quizzes can provide valuable insights for personalized interventions and additional support where needed.

(ii) Considering the crucial role of engagement in student learning, it is recommended to enhance the design of STACK quizzes to maximize positive factors such as feedback, randomization, and peer interaction. Specifically, efforts should focus on improving the clarity of feedback in certain questions to address identified concerns. Additionally, measures should be taken to mitigate challenges related to students' access to technology, ensuring equitable participation. Educators should also explore strategies to boost student motivation, emphasizing the benefits of regular practice with immediate feedback for improved performance.

(iii) To address concerns about the effectiveness of feedback and reluctance to recommend STACK over paper assessments, it is recommended to conduct regular training sessions for both educators and students on maximizing the benefits of STACK. This includes guidance on interpreting and utilizing feedback effectively. Educational institutions should also consider initiatives to improve access to technology for all students. Additionally, continuous communication about the positive impact of STACK on learning outcomes, aligned with the majority preference for continuous assessments, can contribute to a more widespread acceptance of technology-enhanced learning tools in mathematics education.

5.5 Recommendations to the STACK Community.

(i) Foster Collaboration and Knowledge Sharing: Encourage collaboration among educators using the STACK platform to share best practices, effective question formulation techniques, and strategies for promoting student engagement. This collaborative approach can lead to the development of high-quality assessment materials and more impactful learning experiences for students.

(ii) Invest in Professional Development: Provide training and professional development opportunities for educators to enhance their proficiency in using the STACK platform effectively. This can include workshops, webinars, conferences and online resources that focus on maximizing the benefits of technology-based formative assessment.

(iii) Promote Research on Assessment Practices: Support research initiatives that investigate the effectiveness of technology-based formative assessment tools like STACK. Research outcomes can guide the continuous improvement of the platform and inform educators' decision-making in implementing effective assessment strategies.

(iv) Address Accessibility and Equity Concerns: Take steps to ensure that the STACK platform is accessible to all students, including those with disabilities or limited access to technology. Prioritize equity in technology-based assessment to create an inclusive learning environment for all learners. Particularly the offline version of STACK which needs to updated once to increase access beyond internet barriers.

By implementing these recommendations, the STACK community can foster continuous improvement, innovation, and effective use of the platform in formative assessment practices.

5.6 Recommendations for Future Research

(i) Conduct comparative studies between technology-based assessments like STACK and traditional assessment methods to identify the strengths and weaknesses of each approach in terms of student learning outcomes and engagement. This research will provide valuable insights into the effectiveness of different assessment methods and inform decision-making in mathematics education.

(ii) Inclusive Assessment Practices: Examine the accessibility and inclusivity of technology-based formative assessments, including STACK, to ensure that all students, including those with diverse learning needs, can benefit from these tools.

(iii) Investigate the impact of different types of feedback provided by technology-based assessments, such as detailed feedback, on guiding student learning and improving their performance. This research can shed light on the optimal design and delivery of feedback within the context of formative assessment using technology tools like STACK.

(iv) Examine the relationship between the use of technology-based assessments, student motivation, and engagement. Investigate strategies to leverage the benefits of technology tools like STACK to enhance student motivation and promote active engagement with mathematics content.

(v) Investigate the Impact of Different Feedback Types: Conduct further research to explore the impact of various feedback types on cognition in undergraduate mathematics. Understanding the influence of feedback on student learning can lead to more effective instructional practices.

6.0 REFERENCES

Abdool, P. S., Nirula, L., Bonato, S., Rajji, T. K., & Silver, I. L. (2017). Simulation in undergraduate psychiatry: Exploring the depth of learner engagement. *Academic Psychiatry, 41*(2), 251–261.

Abdul Gafoor, K., & Kurukkan, A. (2015). Learner and Teacher Perception on Difficulties in Learning and Teaching Mathematics: Some Implications. In *Online Submission*. Retrieved from https://eric.ed.gov/?id=ED568368

Abramovich, S. (2021). Using Wolfram Alpha with elementary teacher candidates: From more than one correct answer to more than one correct solution. *Mathematics, 9*(17), 2112.

Acee, T. W., Barry, W. J., Flaggs, D. A., Holschuh, J. P., Daniels, S., & Schrauth, M. (2017). Student-perceived interferences to college and mathematics success. *Journal of Developmental Education*, 2–9.

Acosta-Gonzaga, E., & Walet, N. R. (2018). The role of attitudinal factors in mathematical on-line assessments: A study of undergraduate STEM students. *Assessment & Evaluation in Higher Education, 43*(5), 710–726.

Aina, J. K. (2022). STEM Education in Nigeria: Development and Challenges. *Current Research in Language, Literature and Education Vol. 3*, 53–60.

Alabdulaziz, M. S. (2021). COVID-19 and the use of digital technology in mathematics education. *Education and Information Technologies, 26*(6), 7609–7633. https://doi.org/10.1007/s10639-021-10602-3

Al-Hattami, A. A. (2020). E-Assessment of Students Performance During the E-Teaching and Learning. *International Journal of Advanced Science and Technology*, *29*(8), 1537–1547.

Alizadehjamal, M. (2022). Evaluating the performance of Farhangian University mathematics education graduates in terms of technology application based on the model of John Wales and Joseph Bundy. *Technology of Education Journal (TEJ)*.

Alomran, M., & Chia, D. (2018). Automated scoring system for multiple choice test with quick feedback. *Journal of Information and Education Technology*. https://doi.org/10.18178/ijiet.2018.8.8.1096

Alruwais, N., Wills, G., & Wald, M. (2018). Advantages and challenges of using e-assessment. *International Journal of Information and Education Technology*, *8*(1), 34–37.

Ayere, M. A. (2022). Mainstreaming Blended Learning in a Low-Income University. In *E-Learning and Digital Education in the Twenty-First Century*. IntechOpen. https://doi.org/10.5772/intechopen.96143

Bach, Stephan. (2020). *Using multiple choice questions in STACK - reasons and examples*. https://doi.org/10.5281/ZENODO.3945809

Bahati, B., Fors, U., Hansen, P., Nouri, J., & Mukama, E. (2019). Measuring learner satisfaction with formative e-assessment strategies. *International Journal of Emerging Technologies in Learning*, *14*(7).

Baleni, Z. G. (2015). Online formative assessment in higher education: Its pros and cons. *Electronic Journal of E-Learning*, *13*(4), pp228-236.

Barana, A., Marchisio, M., & Rabellino, S. (2019). Empowering engagement through automatic formative assessment. *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*, *1*, 216–225. IEEE.

Barana, A., Marchisio, M., & Sacchet, M. (2021). Interactive Feedback for Learning Mathematics in a Digital Learning Environment. *Education Sciences*, *11*(6), 279. https://doi.org/10.3390/educsci11060279

Bastedo, M. N., Altbach, P. G., & Gumport, P. J. (2023). *American higher education in the twenty-first century: Social, political, and economic challenges*. JHU Press.

Baya'a, N., & Daher, W. (2009). Students' perceptions of Mathematics learning using mobile phones. *Proceedings of the International Conference on Mobile and Computer Aided Learning*, *4*, 1–9.

Beliauskene, E., & Yanuschik, O. (2021). STACK Assessment in Mathematics Classroom: Advantages and Disadvantages. *Educating Engineers for Future Industrial Revolutions: Proceedings of the 23rd International Conference on Interactive Collaborative Learning (ICL2020), Volume 1*, *1*, 174. Springer Nature.

Bethell, G. (2016). *Mathematics Education in Sub-Saharan Africa: Status, Challenges, and Opportunities* [Working Paper]. Washington, DC: World Bank. https://doi.org/10.1596/25289

Bhagat, K. K., & Spector, J. M. (2017). Formative assessment in complex problem-solving domains: The emerging role of assessment technologies. *Journal of Educational Technology & Society*, *20*(4), 312–317.

Biggs, J., & Tang, C. (2007). OUTCOMES-BASED TEACHING AND LEARNING (OBTL). *Why Is It, How Do We Make It Work*.

Blllings, S. A., & Voon, W. S. F. (1986). Correlation based model validity tests for non-linear models. *International Journal of Control*, *44*(1), 235–244.

Blumenfeld, P. C., Kempler, T. M., & Krajcik, J. S. (2006). *Motivation and cognitive engagement in learning environments*. na. Retrieved from https://www.researchgate.net/profile/Toni-Rogat-2/publication/232418824_Motivation_and_Cognitive_Engagement_in_Learning_Environments/links/55c0ca0308aec0e5f447afbb/Motivation-and-Cognitive-Engagement-in-Learning-Environments.pdf

Broekkamp, H., & van Hout-Wolters, B. (2007). The gap between educational research and practice: A literature review, symposium, and questionnaire. *Educational Research and Evaluation*, *13*(3), 203–220.

Bulger, M. E., Mayer, R. E., Almeroth, K. C., & Blau, S. D. (2008). Measuring learner engagement in computer-equipped college classrooms. *Journal of Educational Multimedia and Hypermedia*, *17*(2), 129–143.

Burke, M. J. (1984). Validity generalization: A review and critique of the correlation model. *Personnel Psychology*, *37*(1), 93–115.

Butcher, P. (2008). Online assessment at the Open University us-ing open source software: Moodle, openmark and more. IN: Khandia, F. *12th CAA International Computer Assisted Assessment Conference: Proceed-Ings of the Conference on 8th and 9th July*, 65–78. cc Loughborough University Please cite the published version.

Cardoso, C. (2020). The Challenges Facing African Universities. *Journal of African Cultural Studies*, *32*(3), 302–315. https://doi.org/10.1080/13696815.2019.1671812

Carter-Rau, R., & Olsen, B. (n.d.). Figures of the week: Africa, education, and the 2018 World Development Report. Retrieved November 25, 2023, from Brookings website: https://www.brookings.edu/articles/figures-of-the-week-africa-education-world-development-report-2018/

Chan, Y.-C., & Wong, N.-Y. (2014). Worldviews, religions, and beliefs about teaching and learning: Perception of mathematics teachers with different religious backgrounds. *Educational Studies in Mathematics*, *87*, 251–277.

Chen, B. (2015). Exploring the digital divide: The use of digital technologies in Ontario public schools. *Canadian Journal of Learning and Technology*, *41*(3).

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, *1*, 100002.

Chinofunga, M. D., Chigeza, P., & Taylor, S. (2023). Teachers' perceptions of the effectiveness of a planning framework on content sequencing for the teaching and learning of mathematics. *Eurasia Journal of Mathematics, Science and Technology Education*, *19*(4), em2252.

Codispoti, M., Mazzetti, M., & Bradley, M. M. (2009). Unmasking emotion: Exposure duration and emotional engagement. *Psychophysiology*, *46*(4), 731–738. https://doi.org/10.1111/j.1469-8986.2009.00804.x

Collaço, C. M. (2017). Increasing student engagement in higher education. *Journal of Higher Education Theory and Practice*, *17*(4), 40–47.

Corno, L., & Mandinach, E. B. (1983). The role of cognitive engagement in classroom learning and motivation. *Educational Psychologist*, *18*(2), 88–108. https://doi.org/10.1080/00461528309529266

D'Angelo, C. M., Rutstein, D., & Harris, C. J. (2016). Learning with STEM simulations in the classroom: Findings and trends from a meta-analysis. *Educational Technology*, 58–61.

de Haan, J., Vrancken, J. L. M., & Lukszo, Z. (2011). Why is intelligent technology alone not an intelligent solution? *Futures*, *43*(9), 970–978.

Deng, R., Benckendorff, P., & Gannaway, D. (2020). Learner engagement in MOOCs: Scale development and validation. *British Journal of Educational Technology*, *51*(1), 245–262. https://doi.org/10.1111/bjet.12810

Dermo, J. (2009). e-Assessment and the student learning experience: A survey of student perceptions of e-assessment. *British Journal of Educational Technology*, *40*(2), 203–214.

Derr, K. (2019). optes: Optimising Self-study With STACK. *STACK Online Assessment. A Collection of Case Studies*. Retrieved from https://www.academia.edu/44339885/optes_Optimising_Self_study_With_STACK

Dickinson, K. J., Caldwell, K. E., Graviss, E. A., Nguyen, D. T., Awad, M. M., Tan, S., … Committee, A. E. T. (2021). Assessing learner engagement with virtual educational events: Development of the Virtual In-Class Engagement Measure (VIEM). *The American Journal of Surgery*, *222*(6), 1044–1049.

Dotterer, A. M., & Lowe, K. (2011). Classroom context, school engagement, and academic achievement in early adolescence. *Journal of Youth and Adolescence*, *40*(12), 1649–1660.

Faber, J. M., Luyten, H., & Visscher, A. J. (2017). The effects of a digital formative assessment tool on mathematics achievement and student motivation: Results of a randomized experiment. *Computers & Education*, *106*, 83–96.

Facer, K. (2011). *Learning futures: Education, technology and social change*. Routledge.

Fatima, R. (2012). Role of Mathematics in the Development of Society. *National Meet on Celebration of National Year of Mathematics. Organized by NCERT, New Delhi*, 1–12.

Fynn, A., & Mashile, E. O. (2022). Continuous Online Assessment at a South African Open Distance and e-Learning Institution. *Frontiers in Education*, *7*. Retrieved from https://www.frontiersin.org/article/10.3389/feduc.2022.791271

Gage, M. (2017). Methods of interoperability: Moodle and WeBWorK. *Journal of Learning Analytics*, *4*(2), 22–35.

Gagne, R. (1985). The conditions of learning and theory of instruction Robert Gagné. *New York, NY: Holt, Rinehart Ja Winston*. Retrieved from https://www.ilkogretim-online.org/fulltext/218-1596909338.pdf

Gagné, R. M., & Gagné, R. M. (1985). *Conditions of learning and theory of instruction*. Holt, Rinehart and Winston.

Gammie, E., Hamilton, S., & Gilchrist, V. (2017). Focus group discussions. In *The Routledge companion to qualitative accounting research methods* (pp. 372–386). Routledge.

Gitonga, P., Gatere, R., & Mwaura, J. (2016). *Describing Diversity in University Enrolment in Kenya*.

Goksel, N., & Bozkurt, A. (2019). Artificial intelligence in education: Current insights and future perspectives. In *Handbook of Research on Learning in the Age of Transhumanism* (pp. 224–236). IGI Global.

Goodman, M. S., Sanders Thompson, V. L., Johnson, C. A., Gennarelli, R., Drake, B. F., Bajwa, P., … Bowen, D. (2017). EVALUATING COMMUNITY ENGAGEMENT IN RESEARCH: QUANTITATIVE MEASURE DEVELOPMENT. *Journal of Community Psychology*, *45*(1), 17–32. https://doi.org/10.1002/jcop.21828

Groccia, J. E. (2018). What is student engagement? *New Directions for Teaching and Learning*, *2018*(154), 11–20.

Guangul, F. M., Suhail, A. H., Khalit, M. I., & Khidhir, B. A. (2020). Challenges of remote assessment in higher education in the context of COVID-19: A case study of Middle East College. *Educational Assessment, Evaluation and Accountability*, *32*(4), 519–535. https://doi.org/10.1007/s11092-020-09340-w

Haines, D. (2017). Ethical considerations in qualitative case study research recruiting participants with profound intellectual disabilities. *Research Ethics*, *13*(3–4), 219–232.

Halverson, L. R., & Graham, C. R. (2019). Learner engagement in blended learning environments: A conceptual framework. *Online Learning*, *23*(2), 145–178.

Han, S., & Liou-Mark, J. (2023). Self-efficacy and attitudes towards mathematics of undergraduates: A US and Taiwan comparison. *Journal of Mathematics Education*, 1–15.

Hennink, M. M., Kaiser, B. N., & Marconi, V. C. (2017). Code saturation versus meaning saturation: How many interviews are enough? *Qualitative Health Research*, *27*(4), 591–608.

Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education, 90*, 36–53. https://doi.org/10.1016/j.compedu.2015.09.005

Hodgson, P., & Pang, M. Y. (2012). Effective formative e-assessment of student learning: A study on a statistics course. *Assessment & Evaluation in Higher Education*, *37*(2), 215–225.

Holmes, N. (2015). Student perceptions of their learning and engagement in response to the use of a continuous e-assessment in an undergraduate module. *Assessment & Evaluation in Higher Education*, *40*(1), 1–14.

Innovating Education in Maseno, Kenya—STACK. (n.d.). Retrieved January 31, 2022, from https://stack-assessment.org/CaseStudies/2019/Maseno/

Jamil, M. (2012). Perceptions of University Students regarding Computer Assisted Assessment. *Turkish Online Journal of Educational Technology - TOJET, 11*(3), 267–277.

Jarrah, H. Y., Alwaely, S., & Darawsheh, S. R. (2022). The difficulty and implementation of online assessment in higher education. *Periodicals of Engineering and Natural Sciences*, *10*(1), 692–699.

Jiao, H. (2015). Enhancing students' engagement in learning through a formative e-assessment tool that motivates students to take action on feedback. *Australasian Journal of Engineering Education*, *20*(1), 9–18. https://doi.org/10.7158/D13-002.2015.20.1

Johnson, C. C., Sondergeld, T. A., & Walton, J. B. (2019). A Study of the Implementation of Formative Assessment in Three Large Urban Districts. *American Educational Research Journal*, *56*(6), 2408–2438. https://doi.org/10.3102/0002831219842347

Jones, I. S. (2008). Computer-aided assessment questions in engineering mathematics using MapleTA®. *International Journal of Mathematical Education in Science and Technology*, *39*(3), 341–356.

Kallweit, M. (n.d.). Adaptive Self-learning Exercises. *A Collection of Case Studies*, 17.

Kayode, A. E., & Anwana, E. O. (2023). E-learning of Mathematics and Students' Perceptions in Public Secondary School, Oyo State, Nigeria. In C. Aigbavboa, J. N. Mojekwu, W. D. Thwala, L. Atepor, E. Adinyira, G. Nani, & E. Bamfo-Agyei (Eds.), *Sustainable Education and Development – Sustainable Industrialization and Innovation* (pp. 1077–1087). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-25998-2_83

Keady, G., Fitz-Gerald, G., Gamble, G., & Sangwin, C. (2006). *Computer-aided assessment in mathematical sciences*.

Ketefian, S. (2015). Ethical considerations in research. Focus on vulnerable groups. *Investigación y Educación En Enfermería*, *33*(1), 164–172.

Kim, N. J., Belland, B. R., & Walker, A. E. (2018). Effectiveness of computer-based scaffolding in the context of problem-based learning for STEM education: Bayesian meta-analysis. *Educational Psychology Review*, *30*, 397–429.

King, D. (2023). Assessing the benefits of online formative assessments on student performance. *Journal of Learning Development in Higher Education*. https://doi.org/10.47408/jldhe.vi27.883

Knaut, J., Altieri, M., Bach, S., Strobl, I., & Dechant, K. (2022). A Theory-Based Approach of Feedback in STACK-Based Moodle Quizzes Taking into Account Self-Regulation and Different Proficiency of Learners. *International Journal of Emerging Technologies in Learning*, *17*(23).

Knox, J. (2020). Artificial intelligence and education in China. *Learning, Media and Technology*, *45*(3), 298–311.

Kramer, S., Posner, M. A., Lawrence, N., Browman, A. S., Krier, K., & Roemer, J. (2021, August 31). *Replacing the Traditional Grading System with a System Emphasizing Student-facing Formative Assessment: A Mixed Method Study in Ninth Grade Mathematics Classrooms*. EdArXiv. https://doi.org/10.35542/osf.io/pzc3f

Kramer, S., Posner, M. A., Lawrence, N., Krier, K., Roemer, J., & Browman, A. S. (2021). *Replacing the Traditional Grading System with a System Emphasizing Student-facing Formative Assessment: A Mixed Method Study in Ninth Grade Mathematics Classrooms*.

Lando, H., & Mungan, M. C. (2018). The effect of type-1 error on deterrence. *International Review of Law and Economics*, *53*, 1–8.

Lane, R., Parrila, R., Bower, M., Bull, R., Cavanagh, M., Forbes, A., … Skrebneva, I. (2019). *Literature review: Formative assessment evidence and practice*. Melbourne: Australian Institute for Teaching and School Leadership (AITSL). Retrieved from https://www.lpofai.edu.au/reports/

Lee, J., Park, T., & Davis, R. O. (2018). What affects learner engagement in flipped learning and what predicts its outcomes? *British Journal of Educational Technology*.

Lee, J.-S. (2014). The relationship between student engagement and academic performance: Is it a myth or reality? *The Journal of Educational Research*, *107*(3), 177–185.

Lijie, Z., Zongzhao, M., & Ying, Z. (2020). The influence of mathematics attitude on academic achievement: Intermediary role of mathematics learning engagement. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, *4*(2), 460–467.

Lim, Y. S. (2019). Students' Perception of Formative Assessment as an Instructional Tool in Medical Education. *Medical Science Educator*, *29*(1), 255–263. https://doi.org/10.1007/s40670-018-00687-w

Lowe, M. J., & Vespestad, K. M. (1999). Using technology as a tool to enhance teaching and learning. *NASSP Bulletin*, *83*(607), 30–37.

Lowe, T. W., & Mestel, B. D. (2020). Using STACK to support student learning at masters level: A case study. *Teaching Mathematics and Its Applications: An International Journal of the IMA*, *39*(2), 61–70.

Maamin, M., Maat, S. M., & H. Iksan, Z. (2022). The Influence of Student Engagement on Mathematical Achievement among Secondary School Students. *Mathematics*, *10*(1), 41. https://doi.org/10.3390/math10010041

Maass, K., Geiger, V., Ariza, M. R., & Goos, M. (2019). The role of mathematics in interdisciplinary STEM education. *ZDM*, *51*(6), 869–884.

Mäkelä, A.-M., Ali-Löytty, S., Humaloja, J.-P., Joutsenlahti, J., Kauhanen, J., & Kaarakka, T. (2016). STACK assignments in university mathematics education. *ANNUAL CONFERENCE OF THE EUROPEAN SOCIETY FOR ENGINEERING EDUCATION*. European Society for Engineering Education SEFI.

Mdlalose, N., Ramaila, S., & Ramnarain, U. (2022). Using Kahoot! As A Formative Assessment Tool in Science Teacher Education. *International Journal of Higher Education*, *11*(2).

Modi, K., & Oza, B. (2016). Outlier analysis approaches in data mining. *International Journal of Innovative Research in Technology*, *3*(7), 6–12.

Mugenda, O. M., & Mugenda, A. G. (1999). *Research methods: Quantitative and qualitative approaches*. Acts press.

Mujere, N. (2016). Sampling in research. In *Mixed methods research for improved scientific study* (pp. 107–121). IGI Global.

Nakamura, Y., Ohmata, Y., & Nakahara, T. (2012). Development of a question-authoring tool for math e-learning system stack. *IADIS International Conference E-Learning*, 435–440.

Nakamura, Y., Taniguchi, T., & Nakahara, T. (2014a). Item Bank System for the Mathematics e-Learning System STACK. *Electronic Journal of Mathematics & Technology*, *8*(5).

Nakamura, Y., Taniguchi, T., & Nakahara, T. (2014b). Item Bank System for the Mathematics e-Learning System STACK. *Electronic Journal of Mathematics & Technology*, *8*(5).

Nakamura, Y., Taniguchi, T., & Nakahara, T. (2014c). Item Bank System for the Mathematics e-Learning System STACK. *Electronic Journal of Mathematics & Technology*, *8*(5).

Ngware, M. W., Oketch, M., & Mutisya, M. (2014). Does teaching style explain differences in learner achievement in low and high performing schools in Kenya? *International Journal of Educational Development*, *36*, 3–12.

Nystrand, M., & Gamoran, A. (1991). Instructional discourse, student engagement, and literature achievement. *Research in the Teaching of English*, 261–290.

Ogange, B. O., Agak, J. O., Okelo, K. O., & Kiprotich, P. (2018). Student perceptions of the effectiveness of formative assessment in an online learning environment. *Open Praxis*, *10*(1), 29–39.

Olivier, E., Galand, B., Hospel, V., & Dellisse, S. (2020). Understanding behavioural engagement and achievement: The roles of teaching practices and student sense of competence and task value. *British Journal of Educational Psychology*, *90*(4), 887–909. https://doi.org/10.1111/bjep.12342

Óturai, G., Riener, C., & Martiny, S. E. (2023). Attitudes towards mathematics, achievement, and drop-out intentions among STEM and Non-STEM students in Norway. *International Journal of Educational Research Open*, *4*, 100230.

Oyengo, M. O., Parsons, D., Stern, D., & Sangwin, C. (2021, June). *Providing student feedback through electronic assessment for linear algebra at Maseno University, Kenya*. Presented at the International Meeting of the STACK Community 2021. Zenodo. https://doi.org/10.5281/zenodo.5035969

Özhan, Ş. Ç., & Kocadere, S. A. (2020). The Effects of Flow, Emotional Engagement, and Motivation on Success in a Gamified Online Learning Environment. *Journal of Educational Computing Research*, *57*(8), 2006–2031. https://doi.org/10.1177/0735633118823159

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health*, *42*(5), 533–544. https://doi.org/10.1007/s10488-013-0528-y

Pearson, M. L., Albon, S. P., & Hubball, H. (2015). Case study methodology: Flexibility, rigour, and ethical considerations for the scholarship of teaching and learning. *Canadian Journal for the Scholarship of Teaching and Learning*, *6*(3), 12.

Peytcheva-Forsyth, R., & Aleksieva, L. (2021). Forced introduction of e-assessment during COVID-19 pandemic: How did the students feel about that?(Sofia University case). *AIP Conference Proceedings*, *2333*(1), 050013. AIP Publishing LLC.

Pokharel, T. P. (2023). Students’ Perception of Declining Learner’s Enrollment in Mathematics: A Context of Universities in Nepal. *INTELLIGENCE Journal of Multidisciplinary Research*, *2*(1), 91–102.

Riske, A. K., Cullicott, C. E., Mirzaei, A. M., Jansen, A., & Middleton, J. (2021). Student Engagement with the “Into Math Graph" Tool. *Mathematics Teacher: Learning and Teaching PK-12*, *114*(9), 677–684.

Rowlett, P. (2011). Mathematics lecturers’ views of the advantages and disadvantages of electronic and traditional assessment. *Annual Learning and Teaching Conference Learning and Technology*, 19.

Sangwin, C. (2015a). Computer aided assessment of mathematics using STACK. *Selected Regular Lectures from the 12th International Congress on Mathematical Education*, 695–713. Springer.

Sangwin, C. (2015b). Computer Aided Assessment of Mathematics Using STACK. In S. J. Cho (Ed.), *Selected Regular Lectures from the 12th International Congress on Mathematical Education* (pp. 695–713). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-17187-6_39

Sangwin, C. J. (2002). Encouraging higher level mathematical learning using computer aided assessment. *Trends and Challenges in Mathematics Mathematics Education. Proceedings of the ICM2002 Satellite Conference on Mathematics Education, Lhasa, Tibet*, 255–265. Citeseer.

Sangwin, C. J. (2010). Who uses STACK? A report on the use of the STACK CAA system. *University of Brimingham*.

Schaeffer, C. E., & Konetes, G. D. (2010). Impact of learner engagement on attrition rates and student success in online learning. *International Journal of Instructional Technology & Distance Learning*, *7*(5), 3–9.

Senthilnathan, S. (2019). Usefulness of correlation analysis. *Available at SSRN 3416918*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3416918

Shute, V. j., & Rahimi, S. (2017). Review of computer-based assessment for learning in elementary and secondary education. *Journal of Computer Assisted Learning*, *33*(1), 1–19. https://doi.org/10.1111/jcal.12172

Sikurajapathi, I., Henderson, K. L., & Gwynllyw, R. (2021). Students' Perceptions of Enhanced e-Assessment Feedback Addressing Common Student Errors in Mathematics. *MSOR Connections*, *19*(2).

Smolinsky, L., Marx, B. D., Olafsson, G., & Ma, Y. A. (2020). Computer-based and paper-and-pencil tests: A study in calculus for STEM majors. *Journal of Educational Computing Research*, *58*(7), 1256–1278.

Sun, Y., Guo, Y., & Zhao, Y. (2020). Understanding the determinants of learner engagement in MOOCs: An adaptive structuration perspective. *Computers & Education*, *157*, 103963.

Tomilenko, V. A., & Lazareva, E. G. (2020). Effective Use of a Computer Grading System for Teaching Mathematics at a University. *ITM Web of Conferences*, *35*, 03017. EDP Sciences.

Trowler, P., & Trowler, V. (2010). *Student engagement evidence summary*.

Ullah, A., & Anwar, S. (2020). The effective use of information technology and interactive activities to improve learner engagement. *Education Sciences*, *10*(12), 349.

Ustinova, I., Tomilenko, V., Imas, O., Beliauskene, E., & Yanuschik, O. (2020). STACK Assessment in Mathematics Classroom: Advantages and Disadvantages. *International Conference on Interactive Collaborative Learning*, 174–182. Springer.

VANNATTA, R., BEYERBACH, B., & WALSH, C. (2001). From teaching technology to using technology to enhance student learning: Preservice teachers' changing perceptions of technology infusion. *Journal of Technology and Teacher Education*, *9*(1), 105–127.

Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard university press. Retrieved from https://books.google.com/books?hl=en&lr=&id=RxjjUefze_oC&oi=fnd&pg=PA1&dq=Vygotsky,+L.+S.+(1978).+Mind+in+Society:+The+Development+of+Higher+Psychological+Processes.+Cambridge,+MA:+Harvard+University+Press&ots=ojEZVZq47w&sig=1kNcT-bSbZxQcnS3syaWQgRdsTU

Walker, D. J., Topping, K., & Rodrigues, S. (2008). Student reflections on formative e-assessment: Expectations and perceptions. *Learning, Media and Technology*, *33*(3), 221–234.

Wang, M.-T., Binning, K. R., Del Toro, J., Qin, X., & Zepeda, C. D. (2021). Skill, thrill, and will: The role of metacognition, interest, and self-control in predicting student engagement in mathematics learning over time. *Child Development*, *92*(4), 1369–1387.

Watt, H. M., & Goos, M. (2017). Theoretical foundations of engagement in mathematics. *Mathematics Education Research Journal*, *29*(2), 133–142.

Wilkinson, S. (2011). Analysing focus group data. *Qualitative Research*, *3*, 168–184.

Witkowski, P., & Cornell, T. (2015). An Investigation into Student Engagement in Higher Education Classrooms. *InSight: A Journal of Scholarly Teaching*, *10*, 56–67.

Yang, D., Lavonen, J. M., & Niemi, H. (2018). Online learning engagement: Factors and results-evidence from literature. *Themes in eLearning*, *11*(1), 1–22.

Yang, X. (2013). Investigation of junior secondary students' perceptions of mathematics classroom learning environments in China. *Eurasia Journal of Mathematics, Science and Technology Education*, *9*(3), 273–284.

Zainuddin, Z., Shujahat, M., Haruna, H., & Chu, S. K. W. (2020). The role of gamified e-quizzes on student learning and engagement: An interactive gamification solution for a formative assessment system. *Computers & Education*, *145*, 103729.

Zerva, K. (2020). *STACK: An assessment system for mathematics in Moodle quiz.*

Zerva, K., Sangwin, C., Jones, I., & Quinn, D. (2022). Rejuvenating the HELM Workbooks as Online STACK Quizzes in 2020. *International Journal of Emerging Technologies in Learning (iJET), 17*(23), 69–88. https://doi.org/10.3991/ijet.v17i23.35923

APPENDICES

APPENDIX I: STACK QUIZ ANALYSIS GUIDE

	STUDENT ATTEMPTS ON STACK		DURATION IT HAS TAKEN TO COMPLETE THE QUIZZES		STUDENT SCORE	
WEEK	**MASTERY QUIZ** Frequency of Attempts in the Mastery Quiz	**TEST QUIZ** These Quizzes are restricted to a single attempt and can only be attempted within the stipulated week Done: 1 Not Done: 0	**MASTERY QUIZ**	**TEST QUIZ**	**MASTERY QUIZ SCORE** (Highest Score)	**TEST QUIZ SCORE**
1.						
2.						
3.						
4.						
5.						
6.						
7.						
8.						
9.						
TOTAL SUM					**TOTAL SCORES IN THE MASTERY QUIZZ (A)**	**TOTAL SCORES IN THE TEST QUIZZES (B)**

APPENDIX II: STUDENT'S SCORE CARD AT THE END OF THE SEMESTER

Cat score/ total STACK test score **(Mastery + Tests Quiz) out of 30 marks**	End of semester exam score **Graded out of 30 marks**

APPENDIX III: STUDENT INTERVIEW GUIDE

Introduction

Hello and welcome. Participation in this interview is voluntary and your decision to participate, or not participate, will not affect any services you are currently receiving from the faculty. This is solely for research purposes. This interview should take approximately 20-30min.

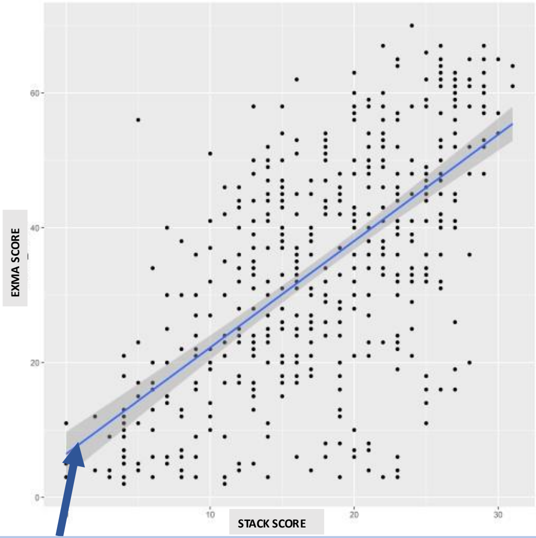
The following scenarios will be investigated using the interview questions presented.

Scenarios	Interview Question for participant.
Scenario 1, where learners show an active engagement with STACK to a point they have relatively high scores in the CAT, and it doesn't reflect on their performance in the exam the following question will be asked.	***These students did well on the STACK Quizzes (They had above average performance in the CAT). However, they did not do well/pass the course as was expected despite having an opportunity to use STACK for practice the entire semester. What might have happened?***
Scenario 2, where learners demonstrate a lack of engagement with STACK, resulting in poor performance on their continuous assessment tests, but perform well on the exam, earning a passing grade in the course. This would contribute to the body of knowledge by explaining how they managed to pull that off despite being provided with the opportunity to use STACK which allows personal practice with immediate feedback.	***These students did well on the Exam, and if only they did better on the STACK Quizzes, they would have passed the course. Why did not they engage with the Quizzes?***
Scenario 3, where learners show not having engaged with STACK resulting in a dismal performance in the continuous assessment tests and the exam which conclusively results in failing the course. This may provide adequate data as to why students are not engaging with STACK and how it should be addressed.	***These are students who passed the course even though they did not do well on STACK. How did they manage to pass the course?***
Scenario 4, where learner engagement with STACK in terms of how they frequently attempt the quizzes is relatively higher than others which in turn doesn't reflect on how they perform at the end of semester examination. This would fill the gap of knowledge on how learners are using feedback to learn and what are the ways of making STACK quizzes effectively helpful in terms of improving on feedback to help them learn.	***Despite having multiple attempts in STACK Quizzes, these students did not seem to learn according to their performance as we would have expected. Can you think of any explanations for this?***

SLIDES USED AS INTERVIEW A GUIDE

The following were used during the Interviews

Graph 1- Relationship between the Exam score and STACK scores



The trend line shows that students who scored higher in the STACK Quizzes performed better on the Exam.

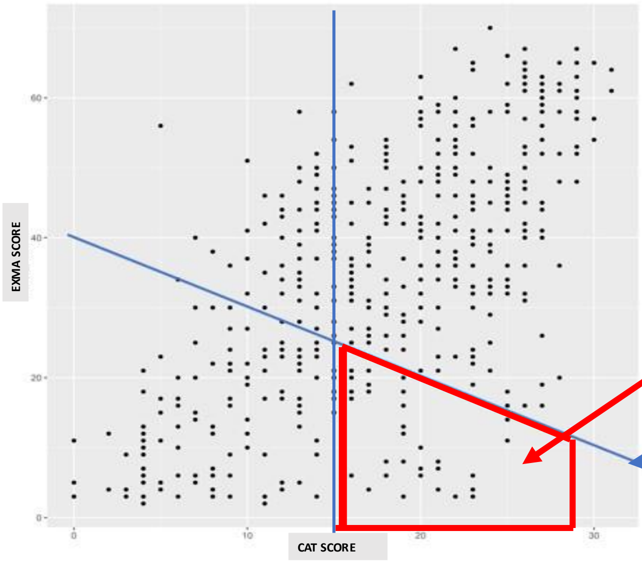
According to the online survey, 97 percent of students felt that the STACK quizzes provided adequate self-assessment tests, while 87 percent felt confident in their problem-solving ability for the exam, attributing it to feedback in STACK. Furthermore, 68 percent stated that they preferred STACK quizzes for Continuous Assessment over non-electronic assessment, such as paper Continuous Assessment Tests.

Question 1
Should we then advise the faculty to integrate STACK in the Continuous Assessment of mathematics courses at Maseno? Why/why not?

Question 2

We've identified 3 groups of students for whom the trend in Graph 1 didn't seem to hold. Can you help us understand what might have happened to each of these sets of students?

Graph 2- Relationship between the Exam score and STACK scores



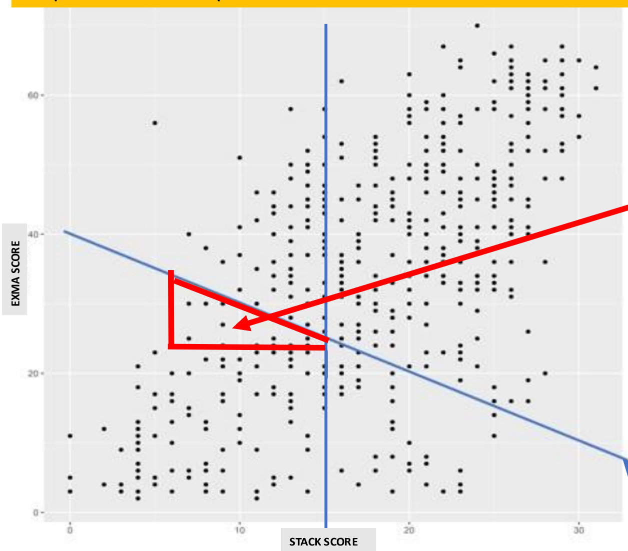
Group A

These students did well on the STACK Quizzes (They had above-average test scores in the CAT). However, they didn't do well/pass the course as was expected despite having an opportunity to use STACK for practice the entire semester. Can you think of any explanation for this?

What might have happened?

Students who passed the Course are above this diagonal line . Those who did not complete the course are identified below.

Graph 2- Relationship between the Exam score and STACK scores



Group B

These students did well on the Exam, and if only they did better on the STACK Quizzes, they would have passed the course.

Why didn't they engage with the Quizzes?

Students who passed the Course are identified above this line. Those who did not pass the course are identified below.

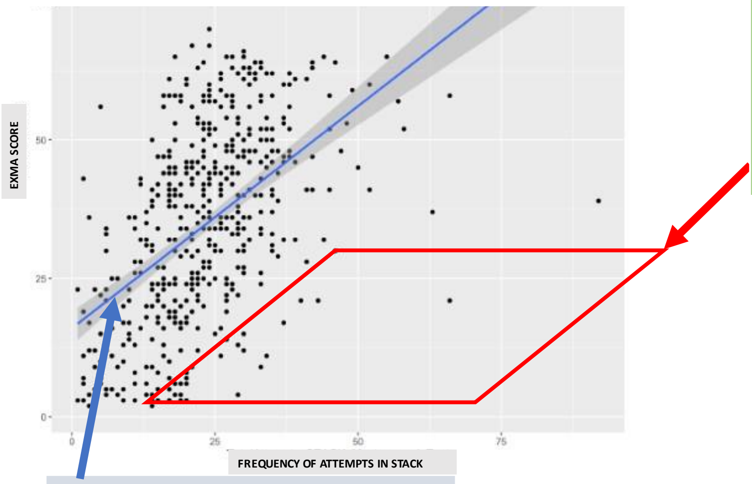
Group C

Graph 2- Relationship between the Exam score and STACK scores

These are students who passed the course even though they didn't do well on STACK.
How did they manage to pass the course yet they didn't engage with the STACK quizzes?

Students who passed the Course are identified above this line. Those who did not complete the course are identified below.

Graph 3- Relationship between the number of STACK attempts and Exam score



Despite having multiple attempts in STACK Quizzes, these students didn't seem to learnaccording to their performanceas we would have expected. Can you think of any explanations for this?

The trend line shows a positive relationship between the number of attempts in STACK and the student's exam score.

APPENDIX IV: SURVEY QUESTIONNAIRE FOR STUDENTS

The purpose of this survey is to collect data on the use of STACK in Continuous Assessment at SMSAS.

Do not provide/ indicate your details anywhere on this survey, we'd wish to maintain anonymity to provide a safe space for learners to express themselves freely.

1. Gender (Select which applies)

MALE		FEMALE	

2. In how many Mathematics courses have you used STACK? ___

3. Which faculty do you belong to?

A. Education
B. Business
C. Computing
D. Science
E. SMSAS

4. Respond to the following comments by selecting the options provided

1-Strongly Disagree, 2-Disagree, 3-Neutral, 4-Agree, 5-Strongly Agree

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
i. I prefer Continuous Assessment Tests with STACK to non-electronic assessments i.e., paper assessment					
ii. There were Adequate self-assessment tests from the STACK quizzes which were enough to prepare me for the Exam					
iii. Having engaged with STACK through practice with the feedback I now feel confident with solving problems in the related content					
iv. Having used STACK for Continuous Assessment, I still would not recommend its adoption over paper assessment in mathematics					
v. I prefer paper assessment to STACK assessment in mathematics at Maseno					
vi. There was enough guidance in the feedback to help me understand the steps outlined in the STACK Quizzes					
vii. STACK doesn't provide adequate self-assessment tests to students while preparing for the end of semester exams					
viii. Using STACK in assessment makes the learning of mathematics more enjoyable and meaningful					
ix. There were instances when feedback was not helping me at all when using STACK					
x. Given an opportunity, I would still not prefer that Continuous assignment given through STACK					

APPENDIX V: FOCUS GROUP DISCUSSION GUIDE

The following slides were used in the Focus Group Discussions

FOCUS GROUP DISCUSSION GUIDE

Introduction

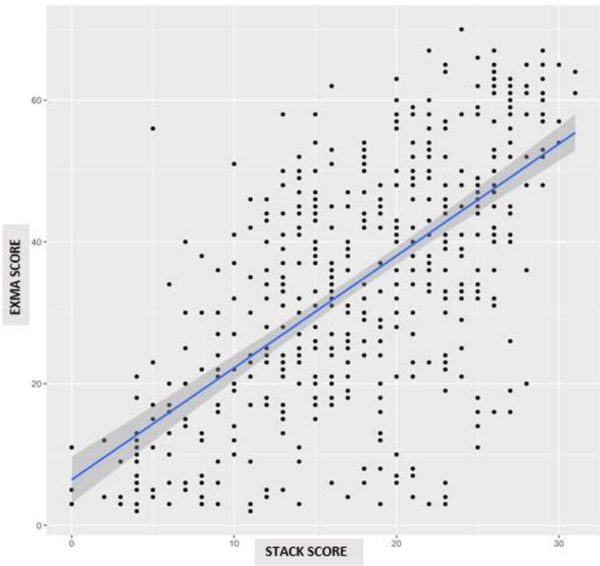
We have one hour for this discussion.

There are 4 slides, each with simple questions intended to stimulate discussion on a specific topic.

On each topic, the facilitator will ensure that each person shares their thoughts and experience.

Differences of opinion are welcome and participants are encouraged to respond respectfully to each other's comments highlighting points of disagreement with explanations justifying their position.

1. Value of STACK: Feedback



87% of students pointed out that **feedback** in STACK gave them the **confidence** to carry on practicing their skills in mathematics and that is why we can see a strong positive correlation between learner performance in STACK and the Exam score.

Furthermore, from the individual interviews, **17 out of 20** students preferred STACK because of the **feedback** they get when they do the assignments.

1. Would you **agree** with these sentiments? If yes, why? If no, why not?
2. What was your personal experience with feedback? How were you using the feedback?
3. Are there any thoughts on how to make the feedback **more helpful?**

2. Potential for STACK: Integration

In individual interviews, **19 out of 20 students** suggested that formative assessment with **STACK** should be implemented more at Maseno and the 20th wasn't against it, but rather had a concern.

In the online survey, **68% of the responses** preferred CATs to be through STACK.

1. **Do you agree**?
2. If so, how should it be implemented? If not, why not?

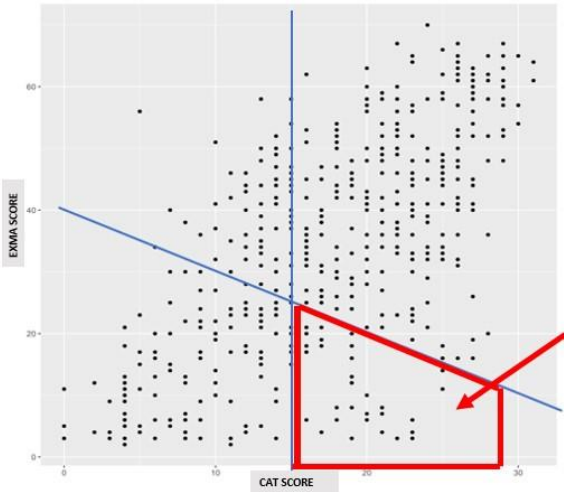
3. Challenge for STACK: Access

According to the responses, students' access to devices and internet connections was identified as a challenge in using STACK. Here is a quote from one of the responses in the interview:

"Some students didn't have smartphones/ internet connections on their phones and therefore they were unable to attempt the quizzes in time…"

1. How did you access the STACK Quizzes? What about the people you know?
2. How much were students with the difficulty of access disadvantaged in the course?
3. What do you think students can do to help each other solve this problem of access?
4. How big is this problem? Can you quantify how many students were affected?
5. Did it affect male and female students differently?

4. Criticism of STACK: Cheating



Some students are clearly **not** learning despite getting high marks in STACK.
In the interviews **17 of the 20 responses** stated that some students who pass the CAT and fail the exam did so because they had others do the assignments for them! Some even have **paid** for the service.

1. Do you agree with this statement?
2. Is this a common practice?
3. Is there resentment amongst other students to this practice?
4. Do you feel anything can/needs to be done to address this? If yes, what is it? If no, why not?

APPENDIX VI: CONSENT FORM
MASENO UNIVERSITY, MAIN CAMPUS

PRIVATE BAG

My name is **Juma Zevick Otieno**, Admission **MED/ED/00027/019**, a Master's student taking a Master’s Degree in Education Technology at the Department of Communication, Pedagogy, Curriculum Studies, and Technology, at Maseno University. I’m doing research on the use of STACK, a computer aided assessment technology being used by lecturers to help with conducting weekly assessment tests in undergraduate mathematics courses at the school of mathematics. I’d wish to engage you in filling an online survey questionnaire, and to also participate in an interview together with a Focus Group Discussion afterwards. Your participation in this will help with providing enough findings that can help future integration of the technology to realize its impact in helping the teaching and learning of undergraduate mathematics.

The following ethical principles will be considered when you choose to participate in this research;

(i) Participation in this study will be entirely voluntary, with no coercion or undue influence, and participant rights, dignity, and autonomy will be respected and appropriately protected.

(ii) No discrimination of any kind will be done to any participant based on ethnic background, level of knowledge, response given amongst others which may arise. Your feedback will be duly respected and recorded as given.

(iii) No names, admission number or any other identifiers of participants will be recorded or shared to the public. All identifiers will be asked to ensure anonymity of responses in this research.

(iv) Other ethical principles protecting participants from any harm related to this research will also be adhered to.

(v) In case one needs more clarification about the project, contact the course lecturer (Dr. Michael Obiero) and me (Juma Zevick).

This is a consent form to acknowledge that you are taking part in this research willingly and without being forced.

Kindly fill out this form if you agree to take part in this research.

I … **(write your initials e.g J.Z)** ………voluntarily agree to participate in this study.

1. I understand that even if I agree to participate now, I can withdraw at any time or refuse to answer any question without any consequences of any kind.

2. I have understood the purpose of this study very well and I feel confident that it is in my interest and that of the body of knowledge to take part in this study

3. I understand that in any report on the results of this research my identity will remain anonymous. This will be done by changing my name and disguising any details I've provided which may reveal my identity or the identity of people I speak about.

4. I understand that I am free to contact any of the people involved in this research i.e. School of Education and School of Mathematics, Statistics, and Actuarial Science to seek further clarification and information.

Participant Signature: ……………. Date: …………………

APPENDIX VII: MAP OF STUDY AREA

Figure 4.1 3.0 Map of Area of Study

APPENDIX VIII: SAMPLE STACK QUESTION USED

Blank Question

Tidy STACK question tool | Question tests & deployed variant

Consider the complex function $f(z) = \frac{\sin\left(\frac{\pi \cdot z}{4}\right)}{(z+1)^3}$.

a) Order of the pole at $z = -1$ is

b) Res$(f(z),\ -1)$ =

Question with worked out solution

Consider the complex function $f(z) = \frac{\sin\left(\frac{\pi \cdot z}{4}\right)}{(z+1)^3}$.

a) Order of the pole at $z = -1$ is 3

Your last answer was interpreted as follows:

$$3$$

✓ Correct answer, well done.

Marks for this submission: 0/0.

b) Res$(f(z),\ -1)$ = %pi^2/2^(11/2)

Your last answer was interpreted as follows:

$$\frac{\pi^2}{2^{\frac{11}{2}}}$$

✓ Correct answer, well done.

Let a complex function be given by $f(z) = \frac{g(z)}{(z-z_0)^{n+1}}$, where $g(z_0) \neq 0$. The function has a pole of order $n+1$ at $z = z_0$. The residue at this pole is given by

$$Res(f(z),\ z_0) = \frac{1}{n!} \lim_{z \to z_0} \frac{d^n}{dz^n} (z - z_0)^{n+1} f(z) = \frac{1}{n!} \lim_{z \to z_0} \frac{d^n}{dz^n} g(z).$$

Comparing with our function $f(z) = \frac{\sin\left(\frac{\pi \cdot z}{4}\right)}{(z+1)^3}$, the pole at $z = -1$ is of order 3 so that $n = 2$. Applying the formula for evaluating residues, the residue at $z = -1$ is given by

$$Res(f(z),\ -1)) = \frac{1}{2!} \lim_{z \to -1} \frac{d^2}{dz^2} (z - -1)^3 \times \frac{\sin\left(\frac{\pi \cdot z}{4}\right)}{(z+1)^3} = \frac{\pi^2}{2^{\frac{11}{2}}}.$$

APPENDIX IX: ETHICS APPROVAL LETTER

MASENO UNIVERSITY SCIENTIFIC AND ETHICS REVIEW COMMITTEE

Tel: +254 057 351 622 Ext: 3050
Fax: +254 057 351 221

Private Bag – 40105, Maseno, Kenya
Email: muerc-secretariate@maseno.ac.ke

REF: MSU/DRPI/MUSERC/01172/22 **Date**: 11th January, 2023

TO: Juma Zevick Otieno
MED/ED/00027/2019
Department of Communication,
Technology & Curriculum Studies
School of Education
Maseno University
P. O. Box, Private Bag, Maseno, Kenya

Dear Sir,

RE: Analysis of the Use and Impact of Systems for Teaching and Assessment which use Computer Algebra Kernel (Stack) in Mathematics at Maseno University, Kenya

This is to inform you that **Maseno University Scientific and Ethics Review Committee (MUSERC)** has reviewed and approved your above research proposal. Your application approval number is MUSERC/01172/22.The approval period is 11th January, 2023 – 10th January, 2024.

This approval is subject to compliance with the following requirements;

i. Only approved documents including (informed consents, study instruments, MTA) will be used.
ii. All changes including (amendments, deviations, and violations) are submitted for review and approval by Maseno University Scientific and Ethics Review Committee (MUSERC).
iii. Death and life threatening problems and serious adverse events or unexpected adverse events whether related or unrelated to the study must be reported to Maseno University Scientific and Ethics Review Committee (MUSERC) within 24 hours of notification.
iv. Any changes, anticipated or otherwise that may increase the risks or affected safety or welfare of study participants and others or affect the integrity of the research must be reported to Maseno University Scientific and Ethics Review Committee (MUSERC) within 24 hours.
v. Clearance for export of biological specimens must be obtained from relevant institutions.
vi. Submission of a request for renewal of approval at least 60 days prior to expiry of the approval period. Attach a comprehensive progress report to support the renewal.
vii. Submission of an executive summary report within 90 days upon completion of the study to Maseno University Scientific and Ethics Review Committee (MUSERC).

Prior to commencing your study, you will be expected to obtain a research license from National Commission for Science, Technology and Innovation (NACOSTI) https://oris.nacosti.go.ke and also obtain other clearances needed.

Yours sincerely

MASENO UNIVERSITY APPROVED 11 JAN 2023 ETHICS REVIEW COMMITTEE

Prof. Philip O. Owuor, PhD, FAAS, FKNAS
Chairman, MUSERC

MASENO UNIVERSITY IS ISO 9001 CERTIFIED