

INTEGRATION OF RUBRICS IN SELF-ASSESSMENT AND PEER-ASSESSMENT: EFFECT ON UNDERGRADUATE STUDENTS' STATISTICS ATTITUDE, SELF- EFFICACY AND PERFORMANCE

AMBROSE AMADIOHA (PhD)

Department of Educational Psychology, Guidance and Counselling
University of Port Harcourt, Port Harcourt, Nigeria

&

AMAKIRI, HAGER ATISI EREMINA (PhD)

Department of Educational Psychology, Guidance and Counselling
Ignatius Ajuru University of Education, Port Harcourt, Rivers State, Nigeria

ABSTRACT

This study was undertaken to investigate the effect of rubrics integration in self-assessment and peer-assessment on attitude, self-efficacy and academic performance of undergraduate students in basic statistics in Rivers State. The non-equivalent quasi-experimental research design was used for the study with a sample of 104 undergraduate students who served as volunteers. The study was guided by three research questions and three corresponding null hypotheses. The study used three major instruments for data collection for three variables under consideration, namely the Survey of Attitudes towards Statistics (SATS), Statistical Self-Efficacy Scale (SSES), and Statistical Performance Test for Undergraduates" (SPTU). Reliability of the SATS and SSES yielded alpha coefficients of 0.85 and 0.90. The SPTU yielded a correlational coefficient of 0.73 using test-retest. Data for the study was collected over eight weeks and involved the collection of both pre-test and post-test scores. Mean, standard deviation and mean difference were used to answer the research questions while analysis of covariance was used to test the null hypotheses were applicable. Results showed that rubrics integrated with peer assessment had the highest positive effect on students' statistical attitude, self-efficacy and academic performance. However, the study also showed that rubrics integrated with self-assessment was more effective than traditional teacher-based assessment. Based on this result, it was recommended that education policymakers should engage more in promoting the use of self and peer assessment in the school system as this has been shown to improve students' academic achievement.

Keywords: Rubrics, authentic assessment, self-assessment, peer-assessment.

Introduction

Assessment has remained an integral part of the formal educational system in the world. While assessment has been considered a fundamental block of modern educational practice, operationalization of the term has been varied due to local and national policies on when and how assessment should be conducted and implemented in the school system. Despite the lack of a universally agreed definition of the term, efforts have been made to provide both theoretical and practical conceptualizations of the term.

A simple definition of assessment by Pellegrino (2006) postulates that assessment refers to the totality of measures used to determine the outcome of education and the achievement of students with regard to identified competencies before and after instruction. The author further proposed that assessment includes both formal and informal methods used at the international, national, state, local and classroom levels with procedures such as quizzes, class projects, and teacher questioning. It is still troubling that the definitions given to assessment are often what determines the assessment approach adopted in schools, which have most times proved defective. This agrees with the opinion of Birenbaum (2003) who asserted that the definition provided for assessment is used as the driving force in the adoption of learning theories, teaching methodologies, competencies and learning goals.

As per the definition provided by Stefanakis as cited by Amakiri and Orluwene (2016), the root word for assessment “assess” is derived from the Latin word “assidene.” The literal meaning of the Latin term is to “sit beside.” On this basis, Stefanakis proposed that assessment literally means to “sit beside the learner to judge.” It was therefore on this basis that the Cambridge Advanced Learner’s Dictionary (2002) stated that assessment originated from the word “assess” which means to judge, decide, or determine the importance, size, or value of something or someone. In this direction, assessment can be stated to mean the process of observing or collecting information on events, objects, or human behaviour to establish the quality or standard of the event, object or person.

In the opinion of Palomba and Banta (1999) assessment involves the systematic collection, review, and use of information about educational programs undertaken to improve student learning and development. In this direction, Huba and Fred (2000) see assessment as the process of gathering and discussing information from multiple and diverse sources to develop a deep understanding of what students know, understand, and can do with their knowledge as a result of their educational experiences; the process culminates when assessment results are used to improve subsequent learning.

Based on the definitions provided above, assessment is defined and used in this study as the process by which information about students' knowledge, skills, ability, belief, aptitude, attitude, performance, competencies and opinions are gathered, analysed, interpreted and utilized to assist students to benefit optimally from a programme of instruction and make decisions on either to continue, modify or terminate the educational programme. Embedded in this definition is the assumption that assessment should not be limited to knowledge and cognitive competencies of students, but rather should involve other affective and psychomotor competencies. Also, this definition recognizes that assessment is not done haphazardly, but involves a systematic and structured process which ought to be followed rigorously to help students optimally. Finally, the definition takes into cognizance the fact that assessment is not an end in itself, but rather serves the purpose of improving students learning, as well as restructuring the learning programme or terminating it if needed, to improve the efficiency of the programme.

In addition to instruction and curriculum development, assessment is another major foundation of any learning experience for students. As such, just as there have been calls for teachers to adopt more effective teaching methods and curriculum planners to modify the curriculum to reflect current realities, there has also been a demand for changes in the

assessment methodologies adopted in contemporary society. Specifically, concerning assessment, modern technologies have provided us with the speed, accuracy and volume of data needed to record, analyze, report and document the learning trajectory and progression of students, but in formal and informal environments. Despite the glaring advantage, research by Anagnostopoulos (2003), as well as a report by Osuji (2015), have argued and showed that the assessment practices of most schools are test-based focusing on the recall of facts and bits of knowledge, with little or no focus on the assessment of complex knowledge and higher-order critical thinking skills.

To forestall the apparent decline of education generally, the adoption of authentic assessment has been proposed (Wiggins, 1990). Cuming and Maxwell (1999) defined authentic assessment as a test of the intellectual achievement or ability of students to demonstrate and apply deep understanding, higher-order thinking, and complex problem solving through the performance of tasks that replicate real-world challenges that students are expected to face. Similarly, Gulikers, et al (2004, p. 69) defined authentic assessment as an assessment requiring students to use the same competencies, or combinations of knowledge, skills, and attitudes that they need to apply in the criterion situation in professional life.

The central aim of authentic assessment is to minimize the artificiality of traditional test-based assessment and propagate the originality of practical, real-world alignment between curriculum objectives, instructional activities and assessment realities. In an attempt to understand and implement authentic assessment, researchers and practitioners have developed various forms of authentic assessments such as non-cognitive assessment, portfolio assessment, project-based assessment, self-assessment, peer-assessment, virtual reality, game-based assessment and the use of rubrics. It is also important to establish that each of these forms of authentic assessment can be combined. It is therefore in this direction that the current study investigates the effect of rubric integration into self-assessment and peer-assessment among undergraduates and the effect on students' attitude, self-efficacy and performance in a basic statistic course.

Self-assessment has been defined as the "involvement of learners in making a judgment about their learning, particularly about their achievement and the outcomes of their learning" (Boud & Falchikov, 1989, p. 529). While students need to be taught how to self-assess themselves, research evidence shows that it is the kind of assessment that students need to effectively develop a deeper understanding of learning contents and themselves. On the other hand, peer assessment occurs when students are required to provide grades on other students' work based on specified and generally agreed-upon criteria. According to Orluwene and Ekim (2020), peer assessment involves a series of processes by which students are given the opportunity of identifying the quality of their peers' work based on an identified benchmark to establish the extent to which the learning objectives have been achieved.

Previous studies have attempted to investigate the impact of self-assessment and peer assessment on attitude, efficacy and performance. For example, Orluwene and Ekim (2020) investigated how students' self-regulated learning can be enhanced through the use of self- and peer-assessment techniques. The result revealed that teacher-assessment, self-assessment, peer-assessment and a combination of self- and peer-assessment techniques were effective in promoting self-regulated learning. In a related study on the impact of self-assessment

training, McDonald and Boud (2010) investigated the impact of self-assessment on the achievement of students. Teachers were selected from a sample of high schools and trained on how to develop students' self-assessment skills. The result revealed a significant difference favouring those trained in self-assessment generally and in each curriculum area. Similarly, Price (2016) investigated the impact of self-assessment on the academic achievement of third-grade students in Mathematics in the United State. The result from the study showed a significant improvement in the academic achievement of the students. From the above-reviewed studies, it is apparent that there are little or no empirical investigations on the impact of rubrics integration in self-assessment and peer-assessment on students' attitude, self-efficacy and performance in statistics in universities in Rivers State.

Aim and Objectives of the Study

The study was aimed at investigating the integration of rubrics into self-assessment and peer-assessment among undergraduates and the effect on students' attitude, self-efficacy and performance in a basic statistic course. In specific terms, this study had the following objectives:

1. To investigate the effect of the integration of rubrics on students' attitude towards statistics using self-and peer-assessment in an undergraduate basic statistics course based on their pre-test and post-test mean scores.
2. To determine the effect of the integration of rubrics on students' self-efficacy in statistics using self-and peer-assessment in an undergraduate basic statistics course based on their pre-test and post-test mean scores.
3. To ascertain the effect of the integration of rubrics on students' performance in statistics using self-and peer-assessment in an undergraduate basic statistics course based on their pre-test and post-test mean scores.

Research Questions

1. What is the effect of the integration of rubrics on students' attitude towards statistics using self-and peer-assessment in an undergraduate basic statistics course based on their pre-test and post-test mean scores?
2. What is the effect of the integration of rubrics on students' self-efficacy in statistics using self-and peer-assessment in an undergraduate basic statistics course based on their pre-test and post-test mean scores?
3. What is the effect of the integration of rubrics on students' performance in statistics using self-and peer-assessment in an undergraduate basic statistics course based on their pre-test and post-test mean scores?

Hypotheses

The following null hypotheses were tested at a 0.05 level of significance to further guide the study:

1. There is no significant effect of the integration of rubrics on students' attitude toward basic statistics using self-and peer-assessment based on their pre-test and post-test mean scores.
2. There is no significant effect of the integration of rubrics on students' self-efficacy in basic statistics using self-and peer assessment based on their pre-test and post-test mean scores.

3. There is no significant effect of the integration of rubrics on students' performance in basic statistics using self-and peer-assessment based on their pre-test and post-test mean scores.

Methodology

Research Design: The non-equivalent quasi-experimental research design was adopted for this study. This design was considered the most appropriate for this investigation because it was done in a school setting involving humans. Akuezilo (2002), stressed that in the school setting, it is not always possible to use a pure experimental design in conducting educational research. The reason is that school authorities may not allow the control or manipulation of some relevant variables which they may consider a disruption of school activities. In this study, two groups of students were assigned to two conditions. While the groups were obtained using randomization, subjects in the groups were not assigned to the groups based on randomization. In the current study, there were two groups, one experimental group and one control group. Both groups received both pre-test and post-test before and after the administration of treatment respectively.

Population of Study: The population of the study was composed of all third-year students in the Department of Business Education, Rivers State University; Department of Educational Psychology, Guidance and Counselling in Ignatius Ajuru University of Education and Department of Educational Psychology, Guidance and Counselling, the University of Port Harcourt during the 2019/2020 academic session. The students were chosen for the study based on the fact that they were taking a course titled "*Basic Statistics and Research Methods*" which was a required course which prepares students for basic and intermediate statistics which can be used in their final research project in their fourth year.

Sample and Sampling Techniques: A non-probability convenience/purposive sampling technique was used for the study. A convenience sampling technique was adopted for the study because it was not possible to randomize the subjects in the current research. Furthermore, this sampling technique was considered appropriate because these researchers were only interested in students who were currently undertaking a course in basic statistics and research methods. Thus, this sampling technique meets the specified need of these researchers. The sample consisted of 104 students who volunteered to participate in an extra-mural class for Basic Statistics and Research Methods. These students were drawn from the departments in the institutions mentioned above. According to the requirements for executing non-equivalent quasi-experimental design, the departments were purposively selected from the three institutions. Students from the three institutions were also randomly assigned to the control group and the experimental groups for the study.

Experimental Procedure: The recruitment of student volunteers was done two weeks into the commencement of the second semester of the 2019/2020 academic session. Those who volunteered were required to fill a form through the Google Form® platform. Initially, a total of 131 students from the three departments volunteered. However, at the commencement of the experiment, only 104 students were present. Students' volunteers were assigned into three groups based on their institution, comprising two experimental groups and one control group. In the first week of training, students were presented with a course syllabus of the various topics that will be covered, the sub-topics, and the needed materials and resources for

participating in the class. Students were also given a schedule of the topics that would be held and the specific times during the work. This was done after ample deliberations to ensure that classes do not clash with the regular class activities of the students. It was also at this point that pretest data for students' attitudes, self-efficacy and performance in statistics were collected.

Students in the self-assessment group were guided on the approach to engaging in self-assessment. This was done by following the suggestion of Andrade and Valtcheva (2009) by creating an awareness of self-assessment, the value and importance of self-awareness, the criteria that will be used for creating self-awareness, as well as the opportunities to revise and improve on task performance. Students were also given the rubrics of the class that would be used in each of the lessons. Grading and feedback are often done immediately after the task performance for students in the self-assessment group.

Students in the peer-assessment group were guided on the approach to engaging in peer assessment. This was done by explaining the value of peer assessment and informing the students of the general procedure that would be adopted. Students were also exposed to the assessment criteria and the opportunities available to them to improve their performance. Students were also given rubrics to guide them in scoring the performance of their peers. In addition to the above, the best practice strategies suggested by Double, et al (2017) were also integrated such as ensuring anonymity in grading by providing random numbers for students to use in their submissions at the end of every week. Students in the peer-assessment groups were encouraged to not only provide a numerical grade but also a written assessment on qualitative feedback. This was to ensure that students thoroughly assess the performance task rather than giving a vague grade which may not represent the effort or performance in the task. Students in this group were given two days to provide grading, feedback and comments on the performance of their peers.

In the control group, students were taught weekly and at the end of the eighth week without any form of assessment or grading. The assessment was only done at the beginning and end of the training period which was used as the pretest and post-test scores. Students were taught using the regularly assigned textbook and grading was done by the researchers.

Instruments for Data Collection: Broadly, three instruments were used for data collection. These instruments were used to assess the variables of statistical attitude, statistical self-efficacy and statistics performance. The first instrument which was used for measuring statistical attitude was titled "Survey of Attitudes towards Statistics (SATS)". This instrument is non-cognitive and was adopted from the instrument developed by Schau (2003). The instrument was composed of 36 items measuring the six components of statistics attitude namely affect, cognitive competence, value, difficulty, interest and effort. The instrument was constructed using a 7-point scale with higher scores corresponding to a more positive attitude. The second instrument for the study was the Statistical Self-Efficacy Scale (SSES). This instrument was an adapted instrument from the Self-Efficacy in Statistical Practice Scale by Pathirage (2015). The researchers adopted the instrument by reducing the number of items from the original 37-items to 15 items. Furthermore, the grading format in the original instrument was done based on the percentage of self-reported confidence. In the current adapted instrument, the researchers used a four-point Likert Scale of Very High Extent (VHE),

High Extent (HE), Low Extent (LE) and Very Low Extent (VLE). This was done because these researchers believe that students are more familiar with it and are more likely to respond. The items were also reduced to minimize the tendency of response fatigue which could result from long items.

The third instrument for this study is a researcher-developed cognitive instrument aimed at assessing students' knowledge of statistical principles and practice. This instrument was titled "Statistical Performance Test for Undergraduates" (SPTU). This instrument was drawn based on the undergraduate syllabus of *Basic Statistics and Research Methods* in the two institutions. The instrument was constructed in an essay format and is composed of five items each focusing on the real-world application of statistics. Students are at liberty to respond to the instrument based on their knowledge and understanding of the statistical principles. Similar instruments were administered to all students in the experimental groups and the control groups. This was done to ensure uniformity of administration and response collection at the pre-test and post-test phases.

Validation of the Instrument: To ensure face and content validities of the instruments were established, copies of the instruments along with the topic of the work, objectives of the study, research questions and hypotheses were given to three experts in Measurement and Evaluation at the Ignatius Ajuru University of Education, Rivers State University, and the University of Port Harcourt. These experts were requested to evaluate the instruments on item content, relevance, comprehensiveness, clarity, and literacy demand. Their suggestions, corrections and modifications would be incorporated into the final version of the instrument before administration.

The reliability of the instruments was established through appropriate statistical techniques. For the SATS and the SSES, the instruments were administered to 20 students in one of the departments not selected for the study. The SATS and SSES yielded Cronbach Alpha coefficients of 0.85 and 0.90 respectively. These values indicate that the instruments had a suitable level of reliability. For the SPTU, the reliability coefficient analysis was done using the test-retest method. To do this, the instrument was administered to the above 20 students twice over a two-week interval. Thereafter, scores from both administrations were correlated using Pearson Product Moment Correlation. The result yielded a correlational coefficient of 0.73 which showed that the instrument possessed a suitable level of temporal stability reliability to be used as a data collection tool.

Administration of the Instrument: Data for the study were collected through the use of the paper-and-pencil format. This format was preferable to the use of an online tool because more students were familiar with it. In addition, the data were collected at the beginning and the end of the study. Data collected at the beginning of the study were referred to as pre-test data while those at the end were referred to as post-test data. In total, eight weeks were used for data collection. The first week was used for the debriefing of the participants and the collection of pretest data. Six weeks were used for instruction and practice of assessment, while the eighth week was used for closure and collection of post-test data.

Method of Data Analysis: In providing answers to the research questions, mean, standard deviation and mean difference were used for both research questions. In testing the null hypotheses, analysis of covariance was used.

Result Presentation

Research Question One: What is the effect of the integration of rubrics on students' attitude towards statistics using self-and peer-assessment in an undergraduate statistics course based on their pre-test and post-test mean scores?

Table 1: Mean and standard deviation of students' statistics attitude in experimental and control groups

Group	n	Mean	SD	Mean Difference
Self Pre-test	35	58.71	6.84	27.78
Self Post-test	35	86.49	4.21	
Peer Pre-test	27	55.89	6.43	42.72
Peer Post-test	27	98.61	5.77	
Control Pretest	42	57.40	5.22	4.77
Control Post-test	42	62.17	6.03	

An observation of the result displayed in Table 1 indicates that when the pretest and post-test values of students' attitude towards statistics in the experimental groups where rubrics were integrated with self-assessment and peer-assessment respectively, as well as those in control, were compared, the mean difference obtained were 27.78, 42.72, and 4.77 respectively. These values indicated that students who were assessed with peer-assessment using rubrics had the highest increase in statistics attitude among the students, followed by those assessed with self-assessment and lastly those in the control groups who were assessed using traditional assessment without rubrics.

Hypothesis One: There is no significant effect of the integration of rubrics on students' attitudes towards statistics using self-assessment and peer-assessment based on their pre-test and post-test mean scores.

Table 2: ANCOVA analysis of the effect of rubrics on students' attitude towards statistics

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	26938.980 ^a	3	8979.660	4632.595	.000	.993
Intercept	1269.661	1	1269.661	655.017	.000	.868
PrStaAt	2765.195	1	2765.195	1426.561	.000	.934
Group	24967.993	2	12483.997	6440.477	.000	.992
Error	193.837	100	1.938			
Total	689657.976	104				
Corrected Total	27132.816	103				

a. R Squared = .993 (Adjusted R Squared = .993)

From the result of the analysis shown above which tests the effect of rubrics integration on students' attitudes towards statistics, the result revealed an F-value of 6440.477 at 2 and 103 degrees of freedom with a corresponding p-value of 0.0005 which was lesser than the chosen alpha guiding the study. Since the p-value was lesser than 0.05 the chosen alpha level, therefore, indicates that integration of rubrics in the self-assessment and peer-assessment had a significant positive effect on students' attitudes towards statistics than the traditional assessment method without rubrics. Furthermore, this yielded a partial eta squared value of 0.992 indicating a strong effect size.

Research Question Two: What is the effect of the integration of rubrics on students' self-efficacy in statistics using self-and peer assessment in an undergraduate statistics course based on their pre-test and post-test mean scores?

Table 3: Mean and standard deviation of students' statistics self-efficacy in experimental and control groups

Group	n	Mean	SD	Mean Difference
Self Pre-test	35	36.28	4.91	25.81
Self Post-test	35	62.09	5.76	
Peer Pre-test	27	38.92	5.04	26.17
Peer Post-test	27	65.09	6.81	
Control Pretest	42	30.71	5.84	9.30
Control Post-test	42	40.01	6.18	

An observation of the result displayed in Table 3 indicates that when the pretest and post-test values of students' statistics self-efficacy in the experimental groups where rubrics was integrated with self-assessment and peer-assessment respectively, as well as those in control, were compared, the mean difference obtained were 25.81, 26.17, and 9.30 respectively. These values indicated that students who were assessed with peer assessment using rubrics had the highest increase in statistics self-efficacy among students, followed by those assessed with self-assessment and lastly those in the control groups who were assessed using traditional assessment without rubrics.

Hypothesis Two: There is no significant effect of the integration of rubrics on students' self-efficacy in statistics using self-and peer assessment based on their pre-test and post-test mean scores.

Table 4: ANCOVA analysis of the effect of rubrics on students' statistics self-efficacy

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	17692.313 ^a	3	5897.438	15215.794	.000	.998
Intercept	497.424	1	497.424	1283.388	.000	.928
PrStaSe	3860.947	1	3860.947	9961.508	.000	.990
Group	4424.601	2	2212.300	5707.887	.000	.991
Error	38.759	100	.388			
Total	320455.312	104				
Corrected Total	17731.071	103				

a. R Squared = .998 (Adjusted R Squared = .998)

From the result of the analysis shown above in Table 4 which tests the effect of rubrics integration on students' statistics self-efficacy, the result revealed an F-value of 5707.887 at 2 and 103 degrees of freedom with a corresponding p-value of 0.0005 which was a lesser than the chosen alpha guiding the study. Since the p-value was lesser than 0.05 the chosen alpha level, it, therefore, indicates that the integration of rubrics in the self-assessment and peer-assessment had a significant positive effect on students' statistics self-efficacy than the traditional assessment method without rubrics. Furthermore, this yielded a partial eta squared value of 0.991 indicating a strong effect size.

Research Question Three: What is the effect of the integration of rubrics on students' performance in statistics using self-and peer-assessment in an undergraduate statistics course based on their pre-test and post-test mean scores?

Table 5: Mean and standard deviation of students' statistics performance in experimental and control groups

Group	n	Mean	SD	Mean Difference
Self Pre-test	35	43.74	8.27	13.36
Self Post-test	35	57.10	7.56	
Peer Pre-test	27	42..98	7.66	18.10
Peer Post-test	27	61.08	8.36	
Control Pretest	42	45.72	6.09	2.99
Control Post-test	42	48.71	8.42	

An observation of the result displayed in Table 5 indicates that when the pretest and post-test values of students' statistics performance in the experimental groups where rubrics was integrated with self-assessment and peer-assessment respectively, as well as those in control, were compared, the mean difference obtained were 13.36, 18.10., and 2.99 respectively. These values indicated that students who were assessed with peer assessment using rubrics had the highest increase in statistics performance among students, followed by those assessed with self-assessment and lastly those in the control groups who were assessed using traditional assessment without rubrics.

Hypothesis Three: There is no significant effect of the integration of rubrics on students' performance in statistics using self-and peer-assessment based on their pre-test and post-test mean scores.

Table 6: ANCOVA analysis of the effect of rubrics on students' statistics performance

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	5351.455 ^a	3	1783.818	1097.025	.000	.971
Intercept	53.259	1	53.259	32.753	.000	.247
PsStaPf	5208.931	1	5208.931	3203.425	.000	.970
Group	2429.872	2	1214.936	747.170	.000	.937
Error	162.605	100	1.626			
Total	21003.046	104				
Corrected Total	5514.060	103				

a. R Squared = .971 (Adjusted R Squared = .970)

From the result of the analysis shown above in Table 6 which tests the effect of rubrics integration on students' statistics self-efficacy, the result revealed an F-value of 747.170 at 2 and 103 degrees of freedom with a corresponding p-value of 0.0005 which was lesser than the chosen alpha guiding the study. Since the p-value was lesser than 0.05 the chosen alpha level, it, therefore, indicates that the integration of rubrics in the self-assessment and peer-assessment had a significant positive effect on students' statistics performance than the traditional assessment method without rubrics. Furthermore, this yielded a partial eta squared value of 0.937 indicating a strong effect size.

Discussion of Findings/ Conclusion

The findings from the study showed that self-assessment had a positive effect on the students' attitude toward statistics. This result was obtained when the pre-test and post-test scores of students were compared after the students were introduced to the rubrics for engaging in self-assessment. The result was expected and not surprising because these researchers believed that allowing students to reflect on the content of a course which is often obtainable through the integration of rubrics facilitates deeper understanding and improved performance. Like most of the cited works in the study corroborated, students' attitude is often a function of their perception. Unless students get to feel the experience of what it takes to engage in statistical thinking, they might have self-defeating thoughts which could hinder their growth. The result from this study is similar to that obtained by other scholars such as Ashari et al (2011) who showed that students' attitude in a course is often a determining factor in the effort they invest in the course.

Similarly, the result reported that students who were exposed to self-assessment had a significantly improved statistics self-efficacy. The findings from this study are in agreement with that of Finney and Schraw (2003) who showed that improved self-efficacy comes from students' understanding of the concepts. The dynamics through which self-assessment promotes self-efficacy might be the opportunity of showing students a task if possible. Rubrics

help to break the task into smaller components as such this might equip students with greater insight and competencies to excel

Finally, the result showed that self-assessment had a significant positive impact on the academic performance of students in statistics. This result might be attributed to the fact that self-assessment shows students the value of self-assessment, access to clear criteria on which to base the assessment, a specific task or performance to assess, models of self-assessment, direct instruction in and assistance with self-assessment, practice, cues regarding when it is appropriate to self-assess, and opportunities to revise and improve the task or performance. The findings of this study are similar to that obtained by Panadero and Jonsson (2013) who established that self-assessment help students easily assimilate the learning objectives and establish more realistic goals, monitor their progress and follow through on improvements. The result further showed that the integration of rubrics in peer assessment was the most effective strategy for improving students' attitudes, self-efficacy and academic performance in statistics. The reason why this may be so might be because peer assessment allows students to critically engage with the assessed material, compare and contrast performance with their peers, and identify gaps or errors in their knowledge. In addition, peer assessment may improve the communication of feedback, as peers may use similar and more accessible language, as well as reduce negative feelings of being evaluated by an authority figure (Liu, et al., 2016). This result of this study is similar to that obtained by Udechukwu (2020) who showed that peer assessment had a significantly greater effect in shaping students' performance in a statistics course than self-assessment. Similarly, Orluwene and Ekim (2020) showed that peer assessment had a significantly higher impact on students' achievement in Chemistry than self-assessment.

Recommendations

Based on the results obtained from the data collection and analysis, the following recommendations were made:

1. Education policy makers should engage more in promoting the use of self and peer assessment in the school system as this has been shown to improve students' academic achievement.
2. Teacher education programs should include programmes and modules to enable teachers to execute self-and peer-assessment effectively in the classroom.
3. Students should be given the opportunity to implement alternative assessment strategies during continuous assessment exercises. This would provide additional opportunities for them to get prompt feedback on their performance against the use of summative assessment at the end of examinations.

References

- Akuezuilo, E. O. (2006), *Research Methodology and statistics*, Owerri; Nulcenti Publishers.
- Amakiri, H. A. E. & Orluwene, G. W. (2016). Differential effect of assessment for learning (AFL) on biology academic achievement of male and female senior secondary school students in Rivers State. *African Journal of Education and Technology*, 6 (1), 1-14.
- Anagnostopoulos D (2003). Testing students' engagement with literature in urban classrooms: A multi-layered perspective. *Research in the Teaching of English* 38(2), 177-192
- Andrade, H. & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory into Practice*, 48(1), 12-19.

- Ashari, N. S., Judi, H. M., Mohamed, H., Tengku, M. & Wook, T. (2011). Student's Attitude towards Statistics Course. *Procedia Social and Behavioral Sciences* 18, 287–294
- Birenbaum, M. (2003). New insights into learning and teaching and the implications for assessment. In M. Segers, F. J. R. C. Dochy, & E. Cascallar (Eds.), *Optimising new modes of assessment: In search of qualities and standards*. Dordrecht: Kluwer Academic Publishers.
- Boud, D., & Falchikov, N. (1989). Quantitative studies of student self-assessment in higher education: A critical analysis of findings. *Higher Education*, 18(5), 529-549. DOI: 10.1007/BF00138746
- Cumming, J. J., & Maxwell, G. S. (1999). Contextualising authentic assessment. *Assessment in Education: Principles, Policy & Practice*, 6(2), 177–194.
- Double, K.S., McGrane, J.A. & Hopfenbeck, T.N. (2017). The Impact of peer assessment on academic performance: A meta-analysis of control group studies. *Education Psychology Review*, 32, 481–509. <https://doi.org/10.1007/s10648-019-09510-3>
- Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development*, 52(3), 67–86.
- Huba, M. E., & Freed, J. E. (2000). *Understanding hallmarks of learner-centred teaching and assessment*. In *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning* (pp. 32-64). Needham Heights, MA: Allyn & Bacon.
- Liu, C., Lu, K., Wu, L. U. & Tsai, C. (2016). The impact of peer review on creative self-efficacy and learning performance in Web 2.0 Learning Activities. *Journal of Educational Technology & Society*, 19(2), 286–297. <http://www.jstor.org/stable/jeductechsoci.19.2.286>
- Orluwene, G. W. & Ekim, D. K. (2020). Promoting self-regulated learning through self-and peer-assessment techniques among secondary school students. *International Journal of Arts and Commerce*, 9(4), 1-16.
- Osuji, U. S. A. (2015). Shifting paradigm in educational assessment. *Global Advanced Research Journal of Education Research and Review*, 4(1), 001-005
- Palomba, C. & Banta, T. W. (1999). *Assessment essentials: Planning, implementing and improving assessment in higher education*: Jossey-Bass, Inc. San Francisco
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited. *Educational Research Review*, 9, 129-144.
- Pathirage, N. A. D. P. (2015). *The development and validation of the self-efficacy in statistical practices scale*. Doctor of Philosophy Thesis submitted to the University of Tennessee-Knoxville.
- Pellegrino, J. W. (2006). *Rethinking and redesigning curriculum, instruction and assessment: what contemporary research and theory suggest*. A Paper Commissioned by the National Center on Education and the Economy for the New Commission on the Skills of the American Workforce
- Price, K. (2016). *The effects of self-assessment on academic performance*. A thesis submitted to Goucher College for a master of Education degree.
- Schau, C. (2003). *Students' attitudes: The "other" important outcome in statistics education*. Paper presented at the Joint Statistics Meetings in San Francisco
- Udechukwu, J. (2020). Effect of self-assessment and peer-assessment techniques on the performance of undergraduate students in a basic statistics course. *International Journal of Innovative Mathematics, Statistics and Energy Policies*, 8(3), 69-75.
- Wiggins, G. (1990). The case for formative assessment. *Practical Assessment, Research and Evaluation*, 2, 8-24.