

Disaggregated Comparative study: Swaziland Junior Certificate 2014 and 2015 history examinations

Boyie S. Dlamini, Sithulisiwe Bhebhe, Phindile A. Dlamini

Department of Education University of Swaziland

Corresponding Author: Boyie S. Dlamini

Abstract: This study examines students' performance in the 2014 and 2015 Junior examination through a disaggregated data lens. Students' exam scores were analysed per exam item, schools and region. This were the questions: what was the nature of the students' performance in the 2014 and 2015 history examinations in Swaziland?; how were the students performing between urban and rural schools?; how were the students performing between the four regions? A sample of 28 schools from 258 was randomly selected. These 28 schools had 1011 students who sat for the 2014 history examination. The total number of students who sat for the examination was 5894. The findings of the study revealed that in both academic years' students' performances in paper one were inadequate. In paper one, of the 2014 and 2015 academic years most schools only performed better in question one. In paper two students' performances were good in question 1, 3, and 4 of 2014 and for the 2015 they performed well only in question 1 and 4. Performances between the regions as well as between urban and rural schools were not clear, because sometimes rural schools outclass urban school across the regions.

Key words: Disaggregated data, Examination scores, Students' performances

Date of Submission: 29-01-2018

Date of acceptance: 15-02-2018

I. INTRODUCTION

Examination scores continue to be used by different governments to measure and judge students' performances and the quality of their education systems but it can provide little insights if they are not analysed in depth (McCallum, 1996; Gurion, 1980). Richer and Brorsen (2006) noted that in the USA thirty five states have comprehensive report systems, reporting test scores and comparisons of schools with district and states but often the information presented is misleading or difficult to interpret. They stated that data arrangement and analysis, which may lead to accurate information on school or students' performance is needed if school reforms are to succeed in improving public schools' performances. Data arrangement involves thinking hard about the data and its implications. For example, focusing on school mean scores when analysing students' performance overlook many factors such as the manner in which students in each school responded to each exam item and how the exam items were chosen, if there were optional questions. Using the students' item mean rather than the school as the only unit of analysis is a way of disaggregating the data and it helps in differentiating students' experiences and performances within schools and among schools and regions. The use of both school mean and students' item mean as the units of analysis is important in examining school and examination effectiveness (Baratz-Snowden, 1993; Mentkowski, 1991), and promoting accountability through examinations. It helps to expose hidden trends and misconceptions about schools and regions' performances.

II. ACCOUNTABILITY THROUGH EXAMINATION

Schools are expected to provide evidence of accountability through examination scores and students' performances. In America, it was recommended that schools should provide evidence of accountability through school assessment and national assessment of educational progress (Downing, 2003). The public and government should hold schools accountable for providing students with knowledge and developing the necessary skills (critical thinking skills; problem solving skills). Examinations are expected to measure the extent to which a student or person commands a certain body of information or skills in the field where training has been received (William et al., 2010; William et al., 2004).

In most education systems, examinations are used as a form of assessment to hold students and schools accountable (Empertson, 1983). The use of examinations as a countable system creates a context in which teachers and policy makers act in ways to maximise performance. Examination based accountability has a

potential of influencing the behaviours of teachers and students in a positive manner (Haertel, 1999; Linn, 1993). Schools are expected to be accountable to parents and government, through the Ministry of Education which is expected to ensure that the national core competencies are addressed in each subject of the examinations. Examinations are expected to show which students and in which schools are meeting the learning standards and which are not (Resnick and Resnick, 1992). Examinations and examination score analysis are central in measuring students' performances per exam item because people need to know about what students have learned as a result of their schooling experiences (Herman, 2005; Anderson, 2002; Baratz-Snowden, 1993). Students' schooling experiences are measured and assessed through every exam item and each examination item score should be analysed in order to identify key concepts and competencies covered and not covered by the examination. This notes that examination assessment is a systematic, connected and purposeful educational process, which focuses on the explicit and implicit links between national competencies and examination items (Hamilton, 2003; Mentkowski, 1991). This requires an interpretive argument (Haertel, 2005; Kane, 1992) which involves obtaining and weighing evidence to support or refute the presented examination reports. An interpretative argument on examination scores has a high educational value because it provides information about students' performances, competencies or on what students have and have not accomplished in regard to specific subject skills and national goals and competencies (Darling-Hammond, 2010; Stecher, et al, 1998).

Relevant educational professionals should ensure that the examination assessment yield information for teachers, parents, and policymakers about what students have learnt, know and are able to do (Wood and Sellers, 1996; Romberg and Wilson, 1992), such information is valid only to the extent that the examination instrument is valid. One of the key indicators of validity of an assessment and examination scores include their alignment with the curriculum objectives (Romberg and Wilson, 1992) and if their analysis and interpretation are supported by appropriate evidence (Messick, 1989; Maduas, 1983). For example, if examination scores are not subjected to deep analysis and the exam items do not reflect the same national competencies goals, subject objectives, content that students are expected to experience in class, then the examination scores cannot be considered a valid means for depicting or providing evidence about students' performances (Romberg and Wilson, 1992). Examination scores should give information about the targeted concepts and students' performances (Messick, 1989). Examination scores provide information on students' capability to show the relationship between the exam items and subject content (Mazer et al. 2008; Williams et al., 2004; Facione, 1989).

III. SWAZILAND JUNIOR CERTIFICATE EXAMINATION

Students who sit for the Junior Secondary (JC) are the ones who have completed the secondary phase of the education system in Swaziland (forms 1-3). Students are examined from a wider range of subjects including history. They are assessed from diversified subjects, which are studied in more depth compared to the Primary Certificate examination.

The JC examination may be considered as the first exit level from the education system because after this level some students may start their independent adult life and enter non-formal education, employment or create their own employment (Ministry of Education and Training Sector Policy, 2011; Ministry of Education Curriculum Framework, 2014). Critical skills such as thinking and communication are consolidated in the education system and examined in an examination. These are important factors for the credibility of the school graduates and the education system (Williams et al., (2010; Downing, 2003; Stecher et al. 2000; Herman, 1997; O' Day and Smith, 1993; Wixson and Pearson, (1989).

IV. STUDENTS' PERFORMANCES

Students' performances are expected to be judged on the bases of the purposes of the test/examination and the examination scores analysis. It should be clearly stated what the examiners want the students to be able to do, and how this requirement fits with the exam instructions, subject content and curriculum goals (Reeves, 2006). The exam should be characterised by an activity which gives the students opportunities to demonstrate the performance and these should align with the subject-matter content and curriculum goal (Reeve, 2006; Brookhart, 1993; Wiggins, 1987). This helps examiners to think of the actual information obtained from students, particularly on how it relates to the intended outcomes (planning and delivering responses-use of knowledge in relevant problem contexts) (Brookhart, 1993).

In several areas of public service, in Swaziland and elsewhere, there is considerable interest in constructing indicators to measure the performance of those services. In education, for example, there are attempts to measure such things such as students' performance and to compare these between institutions in their different settings, rural or urban. This is important in taking educational decisions and in designing targeted educational interventions by policymakers and educational practitioners. Formative analysis of examination scores in relation to students' performances is an evaluation process and a vehicle to improve the quality of subject teaching and educational practices. Evaluation of students' performances is at the centre of educational reforms and to guarantee the quality of education and quality life of the citizens (Bhola et al. 2003).

In Swaziland and elsewhere parents attach more attention and interests on students' examination scores and performance because of their influence on future professional life of the students and the country's economic development. The attractiveness of the examination scores as tools for education policy reflects the importance of not losing sight of students' examination performances. Students' examination performance is linked to the nature of the examination scores and on how they were interpreted for decision making and educational reforms. To guarantee quality educational reforms and quality of education depends on different educational processes such as the validation of the examination scores. Validation of examination scores involves demonstration that the given passing score can be interpreted as representing an appropriate performance standards and the desired level of competence-passing scores and subject domain) (Kane, 1994). This is helpful in taking valid exam score based decisions which depend on the appropriateness of the passing scores used to make the decisions (Kane, 1994). Validity asks whether the proposed interpretation is legitimate, it does not arise until professionals consider interpretations and the proposed interpretation is valid if it is supported by appropriate evidence (Maduas et al, 1993) and an interpretative argument specifying the networks of inferences leading from the score to the conclusions drawn about the students. It involves describing the reasoning involved in the interpreting the performance based on the scores in that particular way (Madaus at al, 1983). These processes are displayed in the analysis of the data or the examination scores. For example, the use of statistical significance analysis is essential because it helps in extracting as much information as possible from the students score and this permits parents, educational practitioner and policymakers to understand interpret and meaningfully discuss published examination results.

In many parts of the world examination scores or standardised testing has become the monitor of students' success and quality of education. For example, legislation, such as the No Child Left Behind Act of 2001 in the United States of America guard schools, on how should be monitored and judged on the basis of students' examination scores performance.

Examination scores determine schooling and job market decisions. They are also considered to signal the ability and cognitive and non-cognitive skills possessed by the person who sits for the examination (O Day and Smith, 1993). This study, therefore, reports validated domain-referenced scores, or scores linked to performance on a domain of items representing the skills and knowledge required for mastery of a content area. The performance is not looked in isolation. Linking examination scores to domain of skills and knowledge shifts focus from individual items to the content domain as a whole. This provides more comprehensive information about students' performance on the 2014 examinations. Though students' academic scores tell just part of the story that goes on in schools, it is one of the recognised measurable outcomes of schools.

V. METHODOLOGY

This is an evaluative study, focusing on the analysis of students' performance scores of the 2014 and 2015 History examination of the Swaziland Junior Certificate. This helped in assessing the examinations' reports credibility. A logical argument to explain the quality of the evidence of students' performance is important and expected in annually examinations reports. Most governments including the government of Swaziland are concerned about their performance in achieving credible exam scores in their public schools and the credibility of exam scores has become a political concern (Page and Feifs, 1985).

Evaluating students' performance from multiple of lenses has the capability to ascertain the validity of the history exam reports on students' performances (Long and Benson, 1989). In this study validity is viewed as a concept with multiple facets, which requires multiple sources of evidence as a form of validation based on validation procedures (Kane, 1992; Downing, 2003 Murphy,1997). Validation of procedures involves the collection of all possible exam item related evidence from different sources to construct an interpretative argument for a convincing report (Resnick et al, 2004; Kane, 1992; Messick, 1989).

This was achieved by paying more attention on those students who performed above the percentage passing mark and below the percentage passing mark score on both the 2014 and 2015 examination. Each examinations consisted of paper one and paper two. Secondly, comparison was made between the questions in each paper by reflecting on difficulty of questions by rank. Thirdly, an ANOVA test for each examination paper was conducted to ascertain whether there is statistical significance between the percentage mean marks for each question.

The evaluation of the students' performance scores during the Junior Certificate examination of 2014 and 2015 was achieved through these following lines of inquiry:

- (a) What was the nature of the students' performance in the 2014 and 2015 History examinations in Swaziland?
- (b) How were the students comparatively performing in the 2014 and 2015 examinations in urban and rural schools in Swaziland?
- (c) How were the students comparatively performing in the 2014 and 2015 History examinations in the four regions of Swaziland?

VI. SAMPLE

A sample of 28 schools from a total of 258 schools which examined students on history examination in 2014 and 2015 was randomly selected. These 28 schools had 1011 students who sat for the 2014 examination. The total number of students who site for the 2014 examination country wide was 5994 (Swaziland Government System Subject Report, 2014). The 1011 (17%) sampled students falls within the required 10% of the 5994 targeted History students population.

The same number of schools (28) was random selected for the 2015 examination. These schools had had 1004 students who site for the 2015 examination. The total number of students who site for the 2015 history examination across the country was 5899 (Swaziland Government System Subject Report, 2015). The 1004 (17%) sampled students falls within the required 10% of the 5899 targeted history students population.

The data used in this study covers a randomly sample of pupils' score in all the four regions of Swaziland. Paper one of the 2014 had a sample of 4009 scores while paper two had 4044 scores. The 2015 examination had a sample of 3986 scores drawn from paper one and 3540 scores from paper two. Paper one had 6 exam items, question **one** and **four** were compulsory while paper two had 4 exam items chosen either from section A or B. Students were expected to answer all the four questions in that chosen section.

VII. RESEARCH DESIGN

Triangulation Design: Convergence and divergence model

The study is anchored on the triangulation design: convergence and divergence model adapted from the Creswell (2003) model. The model is used to compare the students' performances from the 2014 and 2015 results. The ANOVA analysis was conducted through the use of the IBM SPSS Statistics 20 package and this helped in detecting statistical significance and the difficulty level among and between the examination questions (see Table 1; 2). Statistical significance is part of probabilistic inference and helpful in judging whether the scores for students were or not meaningfully different. This helps practitioners not to arrive at unsound conclusion that there was a difference or not, particularly when the pattern in students' performance in different academic years or questions is not completely obvious to all or the performance is ambiguous.

Table 1 Paper 1 2014 History Examination

ANOVA					
Mark	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	683159.579	5	136631.916	337.360	.000
Within Groups	1621226.339	4003	405.003		
Total	2304385.918	4008			

Post Hoc Tests

Homogeneous Subsets

Mark						
Duncan						
Questions	N	Subset for alpha = 0.05				
		1	2	3	4	5
5.00	737	31.2972				
4.00	881		46.7537			
6.00	349			50.2693		
3.00	382				60.9162	
2.00	674					65.3546
1.00	986					66.2535
Sig.		1.000	1.000	1.000	1.000	.450

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 572.905.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Table 2 Paper 2 2014 history Examination

ANOVA

Marks

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	346377.326	3	115459.109	114.439	.000
Within Groups	4076020.091	4040	1008.916		
Total	4422397.417	4043			

Post Hoc Tests

Homogeneous Subsets

Marks

Duncan

Question	N	Subset for alpha = 0.05		
		1	2	3
2.00	1011	43.4451		
3.00	1011		59.2374	
1.00	1011		61.1276	
4.00	1011			68.9357
Sig.		1.000	.181	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 1011.000.

Table 3 Paper 1 2015 History examination

ANOVA

Marks

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	347748.366	5	69549.673	73.860	.000
Within Groups	5700718.425	6054	941.645		
Total	6048466.791	6059			

Homogeneous Subsets

Marks

Duncan

Questions	N	Subset for alpha = 0.05		
		1	2	3
2.00	1005	43.4040		
3.00	1011		59.2374	
4.00	1011		59.2374	
5.00	1011		59.2374	
1.00	1011		61.1276	
6.00	1011			68.9357
Sig.		1.000	.211	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 1009.995.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

Table 4 Paper 2 2015 history examination

ANOVA

Marks

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	346293.134	3	115431.045	114.370	.000
Within Groups	4071408.372	4034	1009.273		
Total	4417701.506	4037			

Homogeneous Subsets

Marks

Duncan

Questions	N	Subset for alpha = 0.05		
		1	2	3
2.00	1005	43.4040		
3.00	1011		59.2374	
1.00	1011		61.1276	
4.00	1011			68.9357
Sig.		1.000	.181	1.000

Means for groups in homogeneous subsets are displayed.
 a. Uses Harmonic Mean Sample Size = 1009.493.
 b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

The model best addresses the research question through different but complementary manner or lens. The model helps by comparing and contrasting the different performances in these academic years of 2014 and 2015. It helps in discovering the paradox and contradictions in students’ performances or results (Greene et al. 1989).

VIII. DATA PRESENTATION

It is important to interrogate the students’ scores in each every question. The ways in which exam scores are used have profound effect not only to the examinees but also on all Citizens because any country’s success and sustainability depends on the attitudes and abilities of the youth to exercise personal judgement as they execute their duties in different societal sectors. Exam scores and education have effects on the development of interpersonal and intrapersonal skills (good judgement and strategies for meeting their own needs in many effective ways) and capabilities that affect the quality of the nation. It is important to disintegrate the students’ performance across the sampled schools for the year 2014 and 2015 and per individual school. This was achieved by paying attention to every student’s performance per exam item.

IX. 2014 AND 2015 PAPER ONE STUDENTS’ PERCENTAGE MARK PERFORMANCE

This paper had six questions, question one and four were compulsory questions. These questions had important concepts and skills expected to be learnt and understood by all students. Question two and three were optional questions and this suggests that the concepts and skills covered in these questions were of less importance compared to those compulsory questions (question one and four). It is important therefore to note how many students responded to each question and how they performed (below the percentage passing mark or above). These would help to shed light on different aspects of the examination questions.

From the sampled schools there were 1011 students who sat for paper one in 2014. In terms of performance, 779 (77%) of the students performed above the percentage passing mark in question one, 218 (22%) of them performed below the percentage passing mark and 14 (1%) did not respond to the question. For the 2015 examination, the sampled schools had 1004 students who sat for the exam. In question one, 478 (48%) of the students performed above passing mark, while 384 (38%) performed below and 20 (2%) did not respond to the question.

In question two, of the 2014 exam 458 (45%) students performed above the percentage passing mark, while 211 (21%) performed below the passing mark and 342 (34%) did not respond to the question. While, in the 2015 exam, 178 (18%) students performed above the passing mark, 115 (11%) performed below and 670 (67%) did not respond.

In question three, of the 2014 exam 231 (23%) performed above the percentage passing mark, while 153 (15%) performed below the passing mark and 626 (62%) did not respond to this question. In the 2015 examination, 313 (31%) of the students performed above the passing mark, 269 (27%) below the mark and 327(33%) did not respond.

In question four of the 2014 examination, a compulsory question, 256 (25%) of students performed above the passing mark, while 636 (63%) of them performed below the passing mark and 117 (12%) did not respond. In the 2015 examination, 281 (28%) of the students performed above the passing mark, while 611 (61%) performed below and 14 (1%) did not respond.

In question five, of the 2014 examination 78 (8%) of the students performed above the percentage passing mark, while 670 (66%) performed below and 262 (26%) did not respond. For the 2015 examination 173 (17%) of the students performed above the passing mark, while 220 (22) performed below and 558 (56%) did not respond.

In question six, of the 2014 examination 133 (13%) performed above the percentage passing mark, while 217 (21%) performed below the passing mark and 660 (65%) did not attempt this question. In the 2015 examination, 223 (22%) performed above the passing mark, while 271 (27%) performed below the mark and 443 (44%) did not respond(see Figure 1: 2014 and 2015 students' performance per exam item).

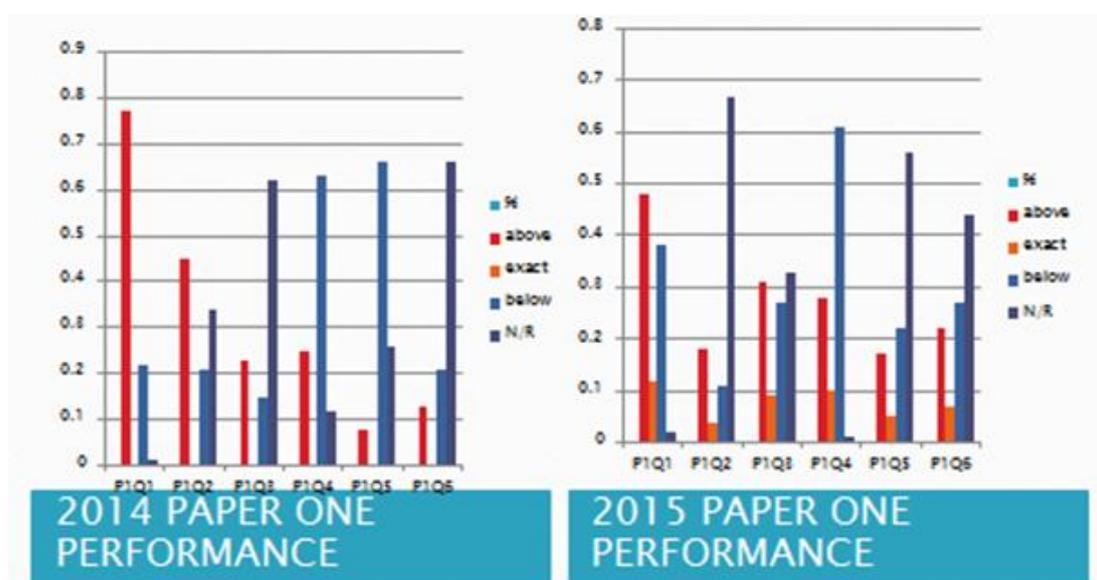


Figure 1 2014 and 2015 students' performance per exam item

From this analysis we observed that in 2014 academic year all the sampled schools' students performed well only in question one, fairly in question two. In 2015 students performed better in question one. In both academic years they performed poorly in question, four, five and six (see Figure 1), and their overall performance was unsatisfactory. In 2014, 22 schools performed poorly (see Table 3). 18 of these 22 schools marked with a star* also performed poorly in 2015 academic year (see Table 3). School FA only performed poorly in 2015 and fairly well in 2014. There were 6 schools which performed better in 2014 and 9 schools in 2015 (see Table 4).

Table 3 Schools performed poorly in 2014 and 2015 (those marked with a star* they performed poorly in both academic years)

Schools	Rural	Urban	Regions
B *		√	Hhohho
LV *	√		Manzini
NH		√	Shiselweni
LB *	√		Hhohho
STA *		√	Manzini
MH *		√	Lubombo
HE		√	Hhohho
EB *		√	Shiselweni
MK		√	Manzini
MD*	√		Shiselweni

MB*	√		Lubombo
KB*		√	Hhohho
MN		√	Manzini
HL*	√		Shiselweni
LS*		√	Lubombo
MS*	√		Manzini
LC*		√	Lubombo
KA*	√		Lubombo
BG*	√		Lubombo
CH*		√	Shiselweni
PH*		√	Manzini
GH*		√	Lubombo

These schools listed in table 3 performed poorly because they had 25% of students or more who performed below the passing mark either in 3 or more exam questions. Overall, in 2014 academic year students performed well in question one, better in two and poor in three, four, five and six. While, in 2015 they performed better in question one and poorly in all other questions (see figure 1)

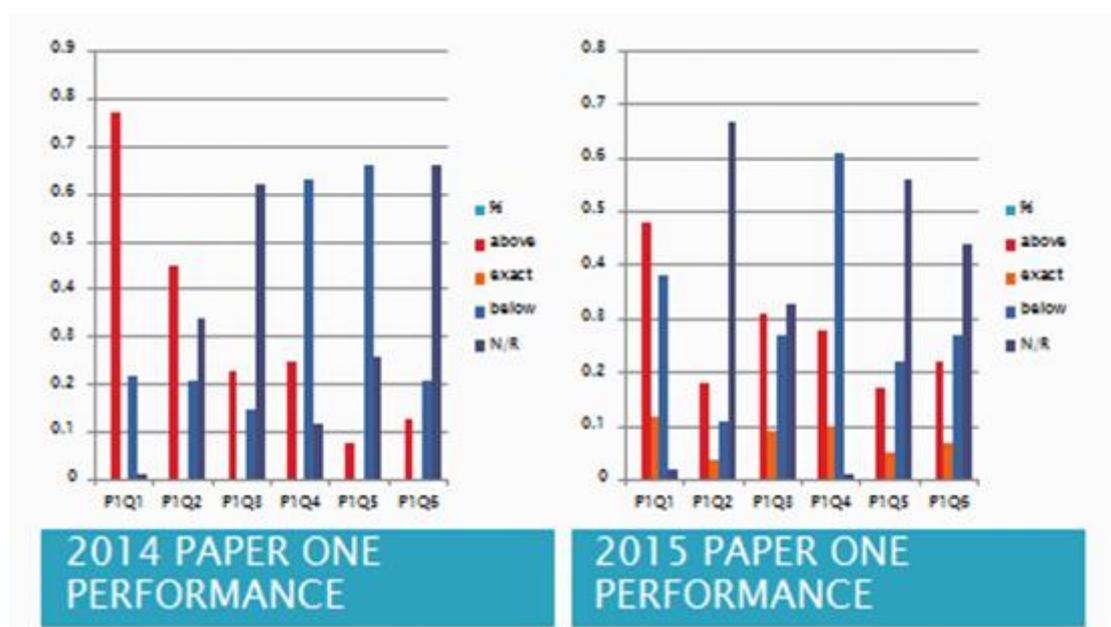


Figure 1; 2014 and 2015 Students' Performance in paper one

Table 4 Schools which performed better in both 2014 and 2015 -paper one

2014 Schools	2015 Schools	Rural	Urban	Regions
MHL	MHL		√	Hhohho
MTS	MTS		√	Manzini
STZ	STZ		√	Manzini
SA	SA		√	Manzini
NY	NY	√		Shiselweni
FA		√		Hhohho
	NH		√	Shiselweni
	MN		√	Manzini
	MK		√	Manzini
	HE		√	Hhohho

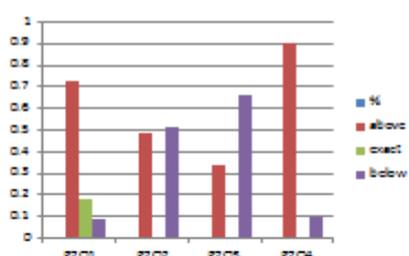
Though there was an increase in the number of schools which performed better in 2015 but only 48% of students performed above the passing mark in question one compared to 77% of students in 2014. As noted earlier, in 2015 students performed better only in question one, while in 2014 they performed well in question one and better in question two. This paper notes that the performance was not improving and this also highlights the importance of focusing on students' performance per question for better understanding of the educational

situation (See Figure 1). The increase of schools as displayed in table 4 does not mean much in terms of meaningful students' performance. In all the questions in 2015, there were less than 50% of students who performed above the passing mark in paper one.

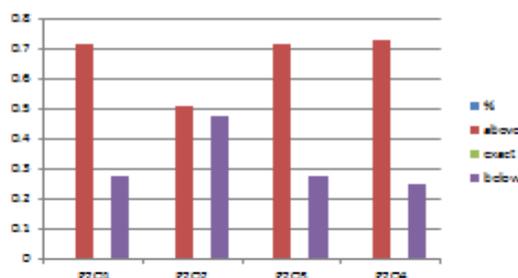
In paper two of 2014, students' performance was good because they performed well in three questions, 1, 3 and 4, while in 2015 they performed well in question 1 and 4 (see Figure 2).

Sampled schools : 2015 and 2014 performance in paper two

2015 paper two



2014 paper two



• Schools performance 2015

• School performance 2014

Figure 2; 2014 and 2015 Students' Performance in paper two

The number of poorly performing schools increased in 2015. There were 14 schools in 2015 and 12 in 2014 (see Table 5). Some of the schools under performed in both 2014 and 2015 academic years (marked with a *).

Table 5 School which poorly performed in 2014 and 2015-paper two

2014 Schools	Sections chosen A or B.	2015 Schools	Sections chosen A or B.	Rural	Urban	Regions
STT	A				√	Manzini
HE	A				√	Hhohho
GH	A				√	Lubombo
MK	A				√	Manzini
MD	B			√		Shiselweni
HL	A				√	Shiselweni
MS*	B	MS*	A		√	Manzini
BG*	B	BG*	B	√		Lubombo
LS*	A	LS*	A		√	Lubombo
PH*	A&B	PH*	A		√	Manzini
MH	A				√	Lubombo
KA*	B	KA*	A&B	√		Lubombo
		MTS	B		√	Manzini
		KB	A&B		√	Hhohho
		SA	A		√	Manzini
		CH	A		√	Shiselweni
		EB	A&B		√	Shiselweni
		STA	A&B		√	Manzini
		BU	A&B		√	Hhohho
		HT	A	√		Shiselweni
		MB	A&B	√		Lubombo

In 2014 section A focused on early European trading activities in Eastern Africa, B on rise of African Nationalism and struggle in Zimbabwe

In 2015 section A focused on Colonisation of Zimbabwe, B on Early European Activities in central and East Africa in 16th to 18th century

Table 5 depicts that most schools performed poorly in both academic years though teachers appeared as if they were teaching for the exam because the majority of schools were either choosing entirely section A or B. This poor performance suggests that teaching for the examination sometimes does not guarantee good performance.

Although more schools performed poorly in 2015 compared to 2014, some of them picked the required sections in a representative manner (see table 5), and this suggests that the concepts were taught.

Table 6 Schools which performed better in both 2014 and 2015 - paper two

2014 Schools	Sections chosen A or B.	2015 Schools	Sections chosen A or B.	Rural	Urban	Regions
B*	A&B	B*	A&B		√	Hhohho
NH*	A&B	NH*	A&B		√	Shiselweni
LV*	A	LV*	A&B	√		Manzini
LB	A&B		A&B	√		Hhohho
MHL*	B	MHL*	A		√	Hhohho
MN*	A	MN*	B		√	Manzini
EB	A		A&B		√	Shiselweni
STZ*	A	STZ*	A&B		√	Manzini
MB	A		A&B	√		lubombo
KB	A		A&B		√	Hhohho
SA	B		A		√	Manzini
LC*	A&B	LC*	A&B		√	Lubombo
FA*	A	FA*	A&B	√		Hhohho
NN*	A	NN*	B	√		Shiselweni
MT	A		B		√	Manzini
	A	GH	A&B		√	Lubombo
	A	MH	A&B		√	Lubombo
	A	MK	A&B		√	Manzini
	A	HE	A&B		√	Hhohho
	A	STA	A&B		√	Manzini

In 2014 section A, focuses on Early European trading activities in Eastern Africa, while B, focuses on rise of African Nationalism and struggle in Zimbabwe

In 2015 section A, focuses on Colonisation of Zimbabwe, while B on Early European Activities in central and East Africa in 16th to 18th century

Table 6 notes that some schools performed better in 2014 and 2015 because teachers were teaching for the examination. For example, school NN, MN, MHL, taught the same concepts drawn from Early European trading activities (see table 6). In 2014 the topic on Early European activities was under section A, while in 2015 it was under section B. In both academic years the students from these schools picked the same topic: Early European activities, though it was placed under different sections (see Table 6). While, other schools such as school B, NH, LV and LC performed better in both years fairly, teachers were not teaching for the examination. Their students picked questions from both sections (A&B), (See table 6).

In the academic year 2015, the 11 schools, (STA, HE, MK, MH, GH, NN, FA, LC, STZ, LV, NH, B) which performed better seemed not to have taught for the examination because their students chose the sections in a representative manner-(A&B) (see table 6), thus guaranteeing that the process of learning had not been undermined.

Through the disaggregated data the study helped in exposing the hidden trends in schools' performances which could not be easily identified through aggregated data. For example, there were 7 (25%) schools in 2014 and 15 (54%) in 2015 which performed unsatisfactorily in question one. In question two, there were 11 (39%) of schools in 2014 and 17 (61%) in 2015. Question three, there were 21 (75) of schools in 2014 and 13 (46%) in 2015. Question four, there were 27 (96%) of schools in 2014 and 14 (50%) in 2015. In question five, there were 26 (93%) of schools in 2014 and 21 (75%) in 2015. In question six, there were 24 (86%) of schools in 2014 and 17 (61%) in 2015 (see table 7).

Figure 7, 2014 and 2015 Schools with unsatisfactory performance per exam item

2015 Schools with unsatisfactory performance per exam item	2014 Schools with unsatisfactory performance per exam item
<ul style="list-style-type: none"> • P1Q1: B,NH,LV,LB, STA,EB,MK,HL,MB,KB,MS,CH,PH,GH, LC, 15 (54%); • P1Q2: B,NH,MHL,EB, STZ,MK,MD,HL,MB,KB,MS,CH,PH, GH,LC,FA, MT, 17 (61%); • P1Q3: B, NH, STA, MD,HL,LS,MS,CH,PH,LC,KA,FA,BG, 13 (46%); • P1Q 4: B,STA, MD,HL,KB,LS,MS,CH,P,GH,LC,KA, BG,MTS, 14 (50%); • P1Q5 : B,NH, STA, MHL, HE, MN, EB, MD,HL,MB,KB,LS,MS,CH,PH,GH,K A,FA,NY,BG,MT, 21 (75%); • P1Q6: LV, LB,STA,MHL,MK,MD,HL,KA,CH,PH ,SA,GH,LC,KA,FA,BG,MTS, 17 (61%). 	<ul style="list-style-type: none"> • P1Q1: MD,CH,PH,GH,LC,KA, BG, 7 (25%); • P1Q2: MHL,MK,MD,HL,CH,PH,GH,LC,KA,BG, MTS, 11 (39%); • P1Q3: B,NH,LV,LB,STA,MHL,HE,STZ,MK,HL,K B,LS,MS,CH,PH,SA, 21 (75%); • P1Q4: B, ,LV,LB,STA,MHL,HE,MN,EB,MN,EB,STZ,MK,MD,HL, MB,KB,LS,MS,CH,PH;SA,GH, LC,KA,FA,NY,MT, BG,27 (96%); • P1Q5: B,NH,LV,STA,MHL,MN,EB,STZ,MK,MD, HL,MB,KB,LS,MS,CH,PH,SA,GH,LC,KA,FA,NY,MT,BG, MTS, 26 (93); • P1Q6: B,NH,LB,STA,MHL,HE,MN,EB,STZ,MK,MD,HL,MB,KB ,LS,MS,CH,PH,SA, GH,LC,KA,FA,N,Y 26 (86%).

Table 7 depicts that most schools performed very poorly in 2014 in question two to six and poorly in question one. Despite the fact that most individual students performed well in question one (see figure 1). This suggests that most of the concepts covered in those questions were missed by the schools, either they were badly taught or not taught. On the other hand, in 2015, all the schools had a very poor performance in all the questions (see table 7).

X. FINDINGS OF THE STUDY

Findings are discussed in relation to the research questions. The research questions are addressed in relation to all the exam items of both papers (paper one and two).

Findings related to research question one: What is the relationship in students' performance in subject skills and knowledge in each exam item for both 2014 and 2015 academic years among the schools in the four regions of Swaziland?

In all the four regions of Swaziland the students performed fairly well in question 1, poorly in question 2 and very poorly in 3, 4, 5 and 6 of the 2014 paper one (see Figure 1-2 and Table 1). Questions such as 3 and 6 were not chosen by many students (Figure 4). This highlights some complexities surrounding these questions. In paper two of 2014, students' performance was good in question 1, 3 and 4 and poorly in question 2 (see Table 4; Figure 1).

The unbalanced representation of the sections in paper two further undermined the students' performances in paper two. The unbalanced representation of the sections suggests that teachers were either teaching for the examination or the students had some problems with the concepts covered in that less chosen section. For the 2015 academic year, students performed better only in question one and poorly in 2, 3, 4, 5 and 6 (see Figure 1, 2 and table 3). In regard to paper one, the study revealed that students had some difficulties with some examination items which require them to interpret the sources and support their interpretation with understanding and knowledge. In paper two of the 2015 academic year, they performed well in question 1 and 4 (see Figure, 2).

Findings related to research question two: Comparison of students' performance between urban and rural schools and regions. 18 schools performed poorly in 2014 and 2015 (see Table 3). 8 were rural schools and 14 urban schools (see Table 3). Urban schools performed poorly compared to rural ones in paper one of 2014 and 2015 (see Table 3). Four schools were from Lubombo region, three from Shiselweni, three from Manzini and two from Hhohho. Lubombo was the worse region among the urban schools. Among the rural schools, three were from Lubombo, two from Manzini, two, from Manzini, two from Hhohho and one from Shiselweni. Again, more poor performing schools were from Lubombo (see Table 3). On the other hand, 15

schools performed better in paper one in 2014 and 2015 (see table 4). There were more urban schools than rural schools and some of these schools performed better in both academic years (see Table 4).

In 2014 there were 6 better performing schools. Manzini had 3, Hhohho had 2, Shiselweni had 1 and 0 from the Lubombo region. In 2015 there were 9 better performing schools (see Table 4), 4 of these 9 schools also performed better in 2014. 5 from Manzini, 2 Hhohho, 2 Shiselweni and 0 in Lubombo. In both academic years urban schools did better in paper one and two (see table 4 and 6). Among the regions, Manzini performed better in both paper one and two of 2014 and 2015, Hhohho region came second position. Whereas Lubombo occupied third position and Shiselweni fourth in paper two. In paper one, of 2014 and 2015 these regions changed positions (Shiselweni third position and Lubombo fourth: see table 4 and 6). This challenges the common misconception that Lubombo is always at the bottom and pointing to the strength of disaggregated data.

Through the disaggregated data the study revealed that many schools' performances were unsatisfactory in question one of both 2014 and 2015. For example, There were 7 (25%) of schools in 2014 and 15 (54%) in 2015 which performed unsatisfactory.

In question two, there were 11 (39%) of schools in 2014 and 17 (61%) in 2015. Question three, there were 21 (75) of schools in 2014 and 13 (46%) in 2015. Question four, there were 27 (96%) of schools in 2014 and 14 (50) in 2015. In question five, there were 26 (93%) of schools in 2014 and 21 (75%) in 2015. In question six, there were 24 (86%) of schools in 2014 and 17 (61%) in 2015 (see table 7).

XI. CONCLUSION

In overall, the students' performances in paper one of both the 2014 and 2015 academic year were unsatisfactory (see figure 1). In paper two of the 2014 examination students performed well in question 1, 3 and 4 and in 2015 they performed better in question 1 and 4 but their performances still remain controversial because it was revealed that there is a possibility that teachers were teaching for examination in some schools (see table 6). It was concluded that only skills and concepts in question **one** and **two** of the 2014 paper one were fairly understood by the students; the rest were either poorly taught or never taught (see Figure 1). In the 2015 paper one, the concepts embraced in question 1, 3 and 4 were fairly understood by the students (see figure 1). It was concluded that in the 2014 and 2015 examinations students found it hard to use instruction words and to interpret historical sources and relate their interpretations to the exam items. The students had a limited ability to connect their interpretations of the sources to what the questions required them to do. In addition, the absence of a relationship or connection between some sources and questions further complicated matters for students (see Source related to question 4 of 2015 and source related to question 4 of 2014: see Appendix 1).

The study concluded that most of the students and schools had difficulties with these instruction words which form a key part of these examinations: answer, analyse, argue, assess, compare, comment, consider, contrast, convey, convince, debate, describe, discuss, explain, explore, interpret, justify, motivate, relate, substantiate.

It was concluded that disaggregated data and its analysis helped in challenging the common misconception that urban schools are performing better than rural schools and that Lubombo is the worst performing region among the four regions of Swaziland. For example, in question five of the 2014 examination 10 schools which performed poorly were from the Manzini region (see Figure 7).

The study also concluded that messages coming from exam scores have a potential of being misunderstood or misused by policymakers and educators, if they are not correctly disaggregated and analysed in details. This involves analysing students' performance per exam item, which could eventually help in giving a fair judgement on students' performances on both urban and rural schools and regions of the country. It is recommended that disaggregated data and detailed analysis should be encouraged by the Ministry of Education and Training because it may expose hidden educational trends and misconceptions on student' performance, and facilitated reforms. It is also recommended that teachers should focus more on instruction words in their teachings so that students can be clear about these words, and be able respond to exam items appropriately.

APPENDIX A

5

SECTION B

Answer question 4 and any other question. Each question is divided into three parts. Answer all parts of the question you choose. Each question carries 15 marks.

Question 4

Study the photograph and answer the questions that follow.



King Sobhuza II with emabutfo

- (a) Describe any one challenge faced by King Sobhuza II when he ascended to the throne in 1921. [3]
- (b) Explain why King Sobhuza II formed the Imbokodvo National Movement in 1964. [5]
- (c) 'Life in Swaziland was better under the colonial rule than after independence.' How far do you agree with the statement? Explain your answer giving both sides of the argument. [7]

Appendix B

SECTION B

Answer question 4 and any other question. Each question is divided into three parts. Answer all parts of the question you choose. Each question carries 15 marks.

QUESTION 4

Study the photograph below and answer the questions that follow.



- (a) Describe any one contribution made by King Mswati III to the development of the Swazi nation. [3]
- (b) Explain why SADC was formed. [5]
- (c) 'Swaziland should not withdraw her SACU membership.' How far do you agree with the statement? Explain your answer giving both sides of the argument. [7]

REFERENCES

- [1]. Anderson, L. W. (2002) Curricular alignment: Re: Examination. *Theory to Practice*, V. 41 (4), Review Bloom's Taxonomy PP 255-260.
- [2]. Baratz-Snowden, J.C. (1993) Opportunity to learn: Implications for Professional Development. *Journal of Negro Education*, v 62, PP 311-323.
- [3]. Bhola, D.S., Impara, J.C., and Buckedahl, C.W. (2003) Aligning tests with States Content Standards: Methodas and Issues. *Educational Measurement: Issues and Practice*, V 22(3), PP 21-29.
- [4]. Brookhart, S.M. (1993) Teachers' Grading Practice: Meaning and values. *Journal of Educational Measurement*. V. 30 (2), PP 123-142.
- [5]. Creswell, W.J (2003) *Research design quantification, quantitative and mixed methods approaches*. Sage: London.

- [6]. Downing, S.M. (2003) Validity: on the meaningful interpretation of assessment data. Department of Medical Education: University of Illinois.
- [7]. Empertson, W.S. (1983) Construct Validity: Construct Representation versus nomothetic Span. *Psychological Bulletin*, V 93 PP 179-197.
- [8]. Faicione, P.H., Sanchez, N.C. and Gainen, J. (1995) The Disposition towards Critically Thinking. *Journal of General Education*. V 44 (1) PP 1-25.
- [9]. Greene, J.C., Coracelli, V.J, and Graham, W.F. (1989) Towards a Conceptual Framework for Mixed Method Evaluation Designs. *Educational Evaluation and Policy Analysis*, V. 11, PP 255-274.
- [10]. Gurion, R.M. (1980) On Trinitarian Conceptions of Validity. *Professional Psychology*, V 11, PP 385-398.
- [11]. Haertel, E. and Herman, J. (2005) a Historical Perspective as Validity Arguments for Accountability Testing. Centre for the Study of Evaluation and Information Studies. University of California. L.A.
- [12]. Darling-Hammond, L. and Adamson, F. (2010) Beyond Basic Skills: The Role of Performance Assessment in Achieving 21st Century Standards in Learning. Stanford Centre for Opportunity Policy in Education. Stanford.
- [13]. Hamilton, L. (2003) Assessment as a Policy Tool. *Review of Research in Education*. V 27, PP 25-68.
- [14]. Herman, J. (1997) Assessment New Assessments: How do they measure up ? *Theory into Practice*, V 36 ,(4), PP 196-204.
- [15]. Kane, M.T. (1992) Validating the Performance Standards Associated with Passing Scores. *Review Educational Research*. V. (3) PP 425-461.
- [16]. Linn, R.L. (1993) Educational Assessment: Expanded Expectations and Challenges. *Educational Evaluation and Policy Analysis*. V. (1), PP 1-16.
- [17]. Long, V.M. and Benson, C. (1989) Alignment. *Mathematics Teacher Journal* V. 19 (6) PP 504-508.
- [18]. McCallum, L. (1996) The Chosen ones? *Education*, V 187, (3), PP 12 -13.
- [19]. Madaus, G.F. and Kellaghan, T. (1993) The British Experience with “Authentic” Testing. *The Phi Delta Kapan*. V 74 (6) pp 458-459.
- [20]. Madaus, G.F.M Stufflebeam, D. and Scriven, M.S. (1983) Programme evaluation: An Historical overview. In G.F. Madaus, M.S. Scriven, and D. Stufflebeam (EDS) *Evaluation*
- [21]. Messick, S. (1989) Evidence and Ethics in the Evolution of tests. *Educational Researcher*. V 10, PP 9-20.
- [22]. Mazer, J.P., Hunt, S.K. and Kuznekoff, J.H. (2008) Reassessing General Education: Assessing a Critical Thinking Instructional Model in the Basic Communication Courses. *Journal of General Education*, V. 56 (3) , PP 173-199.
- [23]. Models: Viewpoints on Educational and Human Services Evaluation (PP 3-22). Pergomon: New York.
- [24]. Mentkowski, M. (1991) Creating a Context where Institutional Assessment Yields Educational Improvement. *Journal of General Education*, V. 40 PP 255-283.
- [25]. Ministry of Education and Training (2011) Ministry of Education and Training Sector Policy 2011. Ministry of Education: Mbabane.
- [26]. Ministry of Education and Training (2014) Swaziland National Curriculum Framework 2014. Ministry of Education: Mbabane.
- [27]. Murphy, R. (1997) Drawing outrageous conclusions from national assessment results: Where will it all end? *British Journal of Curriculum and Assessment*, V7, PP 32-34.
- [28]. O’ Day, J. and Smith, M.S. (1993) Systematic School reform and educational opportunity. In S.H. Fuhrman (ed), *Disingning Coherent education policy: Improving the system*, PP 250-312. San Francisco: Jossey-Bass.
- [29]. Page, E. B. and Feifs, H. (1985) sat Scores and American States: Seeking for Useful Meaning. *Journal of Educational Measurement*, V. 22, (4), PP 305-312.
- [30]. Reeves, T.C. (2006) How do you know they are learning: the importance of alignment in higher education. *Journal of Learning Technology*, V. 2 (4).
- [31]. Resnick, L.B., Rothman, R., Slattery, J.B., Vrennek, J.L. (2004) Benchmarking and Alignment of Students and Testing. *Educational Assessment*. V9 (1, 2), PP 1-27.
- [32]. Resnick, L.B. and Resnick, D.P. (1991) Assessing the Thinking urriculum. *Educational Assessment*. V. (1), PP 36- 65.
- [33]. Richter, F.G.C and Brorsen, B. (2006) Aggregate Versus Disaggregated Data in Measuring School Quality. Oklahoma State University: Oklahoma.
- [34]. Romberg T.A. and Wilson, L.D. (1992) Alignment of Tests with the Standards. *The Arithmetic Teacher*. V. 40, (40), PP 18-22.

- [35]. Stecher, B.M., Barron, S.L., T. and Ross, K. (2000) The Effects of Washington State Education reform on schools and classroom (CSETech. No. 525).Center for Research on evaluation, Standards and students Testing. University of California.
- [36]. Williams, D., Klenowski, V. and Rueda, R. (2010) What's Counts as the Evidence of Educational Assessment? The Role of Constructs in Pursuits of Equity in Assessment. Review of Research in Education. V. 34, PP 254-284.
- [37]. Williams, R.L., Oliver, R. and Stockdale, S. (2004) Psychological Versus Generic Critical Thinking as Predictors and outcome Measures in large undergraduate Human Development Course. Journal of General Education. V. 53 (1) PP 37-58.
- [38]. Wiggins, G. (1993) Assessing students' Performance: Exploring the purpose and limits of testing. San Francisco: Jossey-Bassy.
- [39]. Wixson, K. K., and Pearson, P.D. (1998) Policy and Assessment Strategies to support Literacy Instruction for a New Century. Peabody Journal of Education, V. 73, (3), PP 202-227.
- [40]. Wood, T. and Sellers, P. (1996) Assessment of a Problem –Centred Mathematics: Third Grade. Journal for Research in Mathematics Education. V 27, (3), PP. 337-353.

Boyie S. Dlamini "Disaggregated Comparative study: Swaziland Junior Certificate 2014 and 2015 history examinations." IOSR Journal Of Humanities And Social Science (IOSR-JHSS) , vol. 23, no. 2, 2018, pp. 49-64