Examining Critical and Analytic Thinking in Students' Written Research Projects

Alex Simpande
Copperbelt University
Department of Mathematics and Science Education
P. O. Box 21692, Jambo Drive, Kitwe
Zambia

Leonard Nkhata
Alfred Zulu
Henry Mvula
Copperbelt University
Department of Mathematics and Science Education

Abstract

This study examined critical and analytic thinking in the writing of undergraduate students' research projects. A sample of 23 undergraduate projects was selected for examination. Content Analysis was used to examine the texts (research projects) in detail. A t-test and a two-way ANOVA was further conducted to examine the individual scripts and the six categories of critical and analytic thinking skills. Using the 'discussion of findings' section of the research projects, the analysis addressed the critical and analytic skills of interpretation, analysis, evaluation, inference, explanation, and disposition. The analysis showed that students demonstrate the highest competency in inference and the lowest in disposition. There is no statistically significant difference in students' scores by programme of study and gender. The results highlight aspects of the teaching and acquisition of critical and analytic thinking and how texts portray authors' mastery of these skills.

Key Words: critical thinking, analytic thinking, discussion of findings, content analysis.

1. Introduction

The urgency to reform education and have evidence-based outcomes has characterized the discourse of the Zambian education system, particularly higher education. The role of 21st century skills and competences has taken centre stage in curriculum planning. There are demands placed on both educators and students to be competitive in the global economy and a shift towards incorporating these competences is inevitable. The changes in employment opportunities have accelerated the drive for 21st century skills and competences. The Education for Sustainable Development (ESD) Agenda 2030 and The Africa We Want Agenda 2063 are at the centre of the reform process. Of the 21st century skills, critical and analytic thinking have emerged as cardinal in producing alert graduates in this age of fake news. Kennedy, Fisher, and Ennis (1991) stated that critical thinking is represented by Benjamin Bloom's three highest levels of analysis, synthesis and evaluation. In their revision of Bloom's taxonomy, Anderson and Krathwohl (2001) added *creating* as a higher order thinking skill and it involves generating (hypothesising), planning (designing) and producing (constructing).

Lai (2011) summarizes the critical thinking definitions criteria where most researchers agree: "analyzing arguments, claims, or evidence; making inferences; evaluating; solving problems; asking and answering questions for clarification; identifying assumptions; interpreting and explaining; reasoning verbally and seeing both sides of an issue". Lai (2011) further notes that in addition to these abilities and skills, disposition is an important aspect of critical thinking. These dispositions are referred to as attitudes or habits of mind. Facione (2000) defines critical thinking dispositions as "consistent internal motivations to act toward or respond to persons, events, or circumstances in habitual, yet potentially malleable ways". Although background knowledge or domain specific knowledge is cardinal in critical thinking, researchers have shown that there is more to the process. For example, there is disagreement over the idea of the generality of critical thinking skills and how they can be taught. Following the arguments of Willingham (2007), he states that "it is easier to learn to think critically within a given domain than it is to think critically in a generic sense".

This paper seeks to extend this idea of critical thinking within a given domain by highlighting the application of critical thinking skills in the writing of senior mathematics and science education students.

This approach is contrary to Halpern (2001) and Van Gelder (2005) who uphold the position that critical thinking skills and abilities are not domain-specific. Van Gelder (2005) particularly emphasises that critical thinking content is not tied to any discipline knowledge and only works if taught separately to address issues that students face in their lives.

Kennedy et al. (1991) found that direct instruction through interventions to improve students' critical thinking skills yields positive findings. Abrami et al. (2008), in their analysis of 117 empirical studies 'examining the impact of instructional interventions on students' critical thinking skills and dispositions found this approach to have a positive impact. Changwong, Sukkamart and Sisan (2018) cited a 2015 Thailand Research Fund of 6, 235 students in ten provinces which found that the average score for logical thinking analytical skills was 36.5% with only 2% of all students passing.

Among the many critical thinking measures is the Cornell Critical Thinking Test, (CCTT), the Watson-Glasser Critical Thinking Appraisal (WGCTA-FS), the California Critical Thinking Skills Test (CCTST), among others (Behar-Horenstein, & Niu, 2011).

This study adopted the Saxton, Belanger, and Becker (2012) Critical and Analytic Thinking Rubric (CTAR). The use of rubrics in higher education has become widespread. For example, Campbell (2005) shows that rubrics have been used to assess student work and that they can also serve instructional purposes (Knight, 2006).

The main purpose of this study was to examine the discussion of findings section of final year undergraduate mathematics and science education students' mastery of critical and analytic thinking skills in their projects. The study considered whether scores in these skills differ by programme and gender. The importance of this study is at the core of transforming teacher training in higher education for the 21st century. Its potential to re-orient curricula to ensure the attainment of competence levels is immense. The study was guided by the following research questions:

1. What are the critical and analytic thinking skills demonstrated in the discussion section of research projects?
2. Are there significant differences in performance in critical and analytic thinking skills of students in different programmes?
3. Is there a significant difference in performance in critical and analytic thinking skills between male and female students?

The following are the hypotheses for the study.

1. Research Hypothesis 1: There is a significant difference in critical and analytic thinking skills, by programme of study and gender.
2. Research Hypothesis 2: There is a significant difference in skills performance by programme of study and gender.

2. Methods

This study focused on critical and analytic thinking in undergraduate student-teachers at a public university in Zambia. This is in line with the Ministry of General Education's outcomes based curriculum which places emphasis on 21st century skills and competences (Zambia Curriculum Framework, 2013; Partnership for 21st Century Skills, 2011). The study specifically looked at the discussion section of student-teachers' final year research reports because this section carries the "voice" of a writer and demonstrates the criticalness of the writer.

The units of analysis consisted of 23 student scripts. The study used the content analysis method. Cohen, Manion, and Morrison (2007) state that content analysis is applicable to any written material and also offers room for use to examine large amounts of texts. The data collected are subjected to scientific analysis of meanings and themes that may be both manifest and latent (Cohen, Manion & Morrison, 2007). The units of analysis are students' final year projects, particularly the discussion of findings section of the written projects. The research projects were selected because they reflect a summation of a student's studies in a teacher training program. This approach also gives a cross section of students' knowledge, skills and behaviours related to the skills of critical thinking in writing. In addition, the research projects show how students have mastered (or not) the conventions of critical thinking and communication standards required in writing academic research papers.

Using the R-software, ANOVA tests were used to analyse the data and Bartlett's test of homogeneity of variance was used to test the assumption of constant variance in the one-way ANOVA. A further analysis was conducted using the two-way ANOVA to compare the mean differences in performance scores. The data was also analysed using the Critical Thinking and Analytic Rubric (CTAR) developed by Saxton et.al (2012) to specifically answer research question number one. The rubric consists of six main areas namely: interpretation, analysis, evaluation, inference, explanation and disposition. Each of these is measured on a scale of 1 to 6 with 1 being the lowest and 6 being the highest score. Descriptors are indicated at each level of the score.

The selection of projects in focus was not based on an ESD criteria but were selected to identify how students demonstrate the ESD 21st century skills and competences of critical and analytic thinking. The subject offerings in the education department at the site are physics, biology, chemistry and mathematics. Only physics, biology and mathematics projects were selected. Simple random sampling was used to select the projects from each subject area list. Nine projects were selected from physics and nine from biology. However, only five were selected from mathematics and none from chemistry. The reduction in mathematics project scripts and the exclusion criteria for chemistry project scripts was solely based on the time frame required to complete the study, which in itself is a limitation.

3. Results

In order to realise the goals of this study, qualitative data in the form of rubric scores analysis was done while quantitative data in the form of descriptive statistics were used to show the means and visual representation of the data. The rubric analysis provided an in-depth examination analysis of the actual qualitative meaning by depicting the relationships between the skills, programmes and gender and also the implications for professional practice.

3.1 Research Question One: What are the critical and analytic thinking skills demonstrated by mathematics and science education students in the discussion section of their research projects?

3.1.1 Rubric Analysis

The Critical Thinking Analytic Rubric (CTAR) developed by Saxton, Belanger and Becker (2012) was used to score the scripts against the descriptors in the rubric. Saxton et al. (2012) used this rubric to analyse the writing of 304 high school students in the initial study. The rubric is grounded in the Delphi definition of critical thinking and the analytic rubric gives important data for critical thinking and studies related to writing for writers in higher education. The rubric contains six categories of skills: interpretation, analysis, evaluation, inference, explanation and disposition. Each category has descriptors scored between 1 and 6, with 1 being the lowest and 6 the highest. The highest (total) score that a script can attain for all the six scores is 36 and the lowest score is 6. The total maximum score per skill for all 23 scripts is 138.

The average overall scores from both raters for each skill show that inference, with a score of 46, is the highest score followed by analysis with 45, then evaluation with 42. Interpretation follows with 39.5, explanation with 37.5 and disposition with 37.

Table 1 shows a summary of the aggregated scores of both raters. The results show that the rating was fairly consistent except in two instances that involved script 7 and script 13. Consistency here is defined as a difference of less than six between raters since that is the highest score for any skill and the lowest skill that a script could score.

Table 1. Rating results per script for all six skills

	Rater 1 (Overall Score)	**Rater 2** (Overall Score)
Script 1	11	11
Script 2	12	8
Script 3	7	8
Script 4	7	6
Script 5	17	13
Script 6	6	9
Script 7	22	8
Script 8	13	14
Script 9	8	11
Script 10	9	7
Script 11	9	8
Script 12	13	8
Script 13	17	6
Script 14	8	10
Script 15	22	20
Script 16	10	12
Script 17	9	9

Script 18	8	9
Script 19	9	12
Script 20	13	13
Script 21	9	10
Script 22	12	11
Script 23	19	16

The two instances where the difference was beyond six are highlighted in the table. This is in relation to script 7 which had a difference of fourteen and script 13 with a difference of eleven between the raters. A third rater was involved to grade script 7 and script 13 using the same rubric. The results are displayed in Table 2.

Table 2. Third rater scores of scripts 7 and 13

	Program	**Sex**	**Interpr.**	**Analy.**	**Eval.**	**Infer.**	**Explan.**	**Disp.**	**Total**
Script 7	P	M	3	2	2	2	1	1	**11**
Script 13	P	M	2	2	1	3	2	1	**11**

A content analysis of the scripts using the Saxton et al. (2012) rubric reveals certain patterns in the discussion section of the research report. The manifest patterns evident across the skills are consistent and bring into focus the conventions of research report writing for the sampled mathematics and science education students. A summary of the manifest patterns in the scripts is provided below.

3.1.1.1 Interpretation

Interpretation is evident in being able to clearly and accurately identify all the major viewpoints; interpret evidence, statements, graphics and questions; being able to work with key concepts and terminology. The results show that descriptor three is the least scored, descriptor two is the highest with six scores. No author scored between four and six. The writer fails to connect her findings to other studies and does not connect the literature review with her findings**.** This inability to connect findings to other studies is evident in other scripts namely 2, 4, 5, 6, 7, 8, 9, 11.

3.1.1.2 Analysis

Performance in the skills of analysis, evaluation and inference show that raters were fairly consistent in the awarding of scores and the variation was not significant. The *t* statistic in table 3 demonstrates the means between the raters. In terms of the skill of analysis, results show that the highest rubric score was four and it occurred once and the rubric score of three and one were recorded once each. The most frequent score under analysis was the rubric score of two. The majority of scripts presents a superficial analysis of similarities and differences between the various points of view; incorrectly identifies claims, arguments, patterns, and/or assumptions in the evidence and demonstrates an inadequate ability to organize the information for further examination. For example, in descriptor two, script two, the writer lists the challenges experienced by the pupils *"identified by the researcher"*. However, the researcher does not indicate how these challenges were collected.

3.1.1.3 Evaluation

Evaluation is defined as being able to identify salient arguments (reasons and claims) from multiple perspectives with a clear explanation of each perspective; thoughtfully analyzing all major alternative points of view/counter-arguments. While one student author is able to identify relevant arguments from multiple perspectives and offer analyses and evaluations of alternative points of view, the majority of student authors in this category superficially identifies some arguments (reasons and claims) from main perspectives and superficially evaluates obvious alternative points of view. For example, descriptor one, in script five, there are no specific arguments or counter-arguments provided. In descriptor two of script nine, there is no connection of the results to the literature review or any study done on this topic. This leads to a lack of any other argument or point of view to contextualize the evidence and claims.

3.1.1.4 Inference

Inference is the ability to apply or extend key concepts to make predictions, drawing conjectures and analyzing implications and demonstrate surprising or insightful ability to take concepts further into new territory with broader generalizations and implications. In descriptor three, script ten, there is inadequate interpretation of the teacher questionnaire which has affected the activity of applying and extending the key concepts to analyze the implications.

3.1.1.5 Explanation

Although the overall scores show that there is no statistically significant difference in students' performance in the six skills, there is a variation in the awarding of scores between rater one and rater two as shown in the box plots. Rater two has lower scores for the skills of explanation (evidence) and interpretation although this difference is not significant as reflected in figure 1 and figure 2. *Explanation* means integrating sources to support conclusions, justify/explain assumptions and reason with evidence (references). As stated in the theme of analysis, some scripts such as 1, 2, 4, 5, 6, 7, 8, 9, and 11 do not adequately reference information in the discussion section. The inability to relate author findings to the literature review, counter-arguments and other studies occurs in scripts *2, 4, 5, 6, 7, 8, 9, 10* and descriptor one and two, the lowest descriptors on the rubric, apply to these scripts.

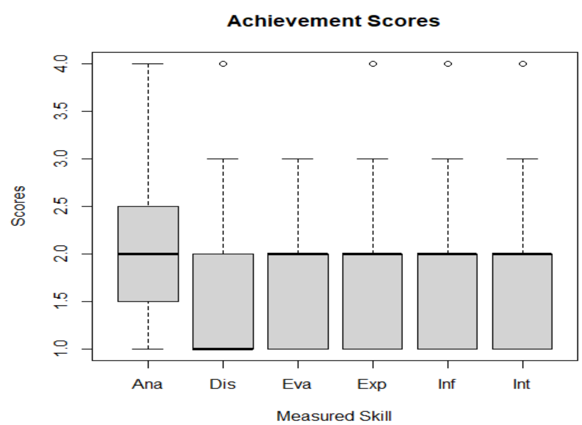
3.1.1.6 Disposition

The results show that the skill in which students demonstrate less competence is disposition. There is consistency in the raters' awarding of scores for this skill and only a slight variation in the skills of explanation and interpretation as shown in figure 1 and figure 2. Saxton et al. (2012) state that disposition refers "follow where the evidence leads by considering the provided context; demonstrate relativist view of knowledge through adoption of a consistent point of view with appropriate justification and awareness of alternative viewpoints. Soper (2015) adapted Saxton's rubric and defined disposition as "the writer shows empathy and open-mindedness for other viewpoints, and displays an ethos of diligence, care, reasonableness, and persistence in attempts to think critically". For example, descriptor one, scripts 7, 8, 9, and 10 demonstrate close-mindedness or hostility to reason and focus on one side of the issue and do not display an ethos of diligence. All the four scripts show actual images of students (sometimes the names) with no consent form.

In summary, when the results of the six skills are considered together, the evidence shows that the skill of inference is more pronounced than all other skills while disposition is the skill where students demonstrate less competence.

The box plots below demonstrate the occurrence of the six skills as scored by each rater.

Figure 1.Box plot Rater 1



Note. Representation of the distribution of achievement scores data

The horizontal 'dark' line in the boxes indicates the performance scores. Thus it can be seen that there are no significant differences in scores within skills, except for the "Disposition" skill which seem to have a different mean.

Figure 2.Box plot Rater 2

Note. Representation of the distribution of achievement scores data

Evidently for rater 2, there are noticeable differences in the performance scores from the plot, when the horizontal dark lines are considered.

Table 3. Summary statistics for the *t* distribution:

Raters	Replicates (n_i)	Mean ($\bar{y}$)	Variance (s^2)	SD (S_y)	t	P-value
Rater1	23	11.74	22.02	4.49	1.127	0.266
Rater2	23	10.39	10.98	3.31		

Adopting a 0.01 conventional significant level, we fail to reject the null hypothesis of no difference between the overall scores of raters 1 and 2. Therefore we can conclude that the raters were fairly consistent in the awarding of scores across the six categories.

3.2 Research question two: Are there significant differences in critical and analytic thinking skills of students in different programmes?

The rubric analysis by programme is indicated in Table 4. The number of scripts for Physics and Biology were nine each while mathematics scripts were five. It was important to answer the question about differences in skills by programme to determine whether the prevalence of the six skills in the rubric vary by programme which would suggest a particular emphasis in each programme.

The total score for each skill in the programme was computed and the maximum possible score for all the skills combined is 216. For rater 1, the total scores for Physics were 115, Biology 111 and Mathematics 44. The total score in mathematics is much lower than Physics and Biology because only five scripts were analysed compared to the other two programmes which had nine scripts each.

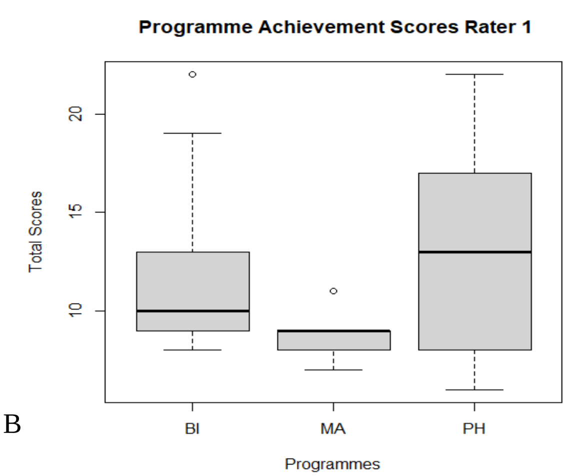
3.2.1 Rater 1 Scores by programme

Table 4: Rater 1 one-way ANOVA analysis by programme

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F	Pr (>F)
Programmes	2	56.1	28.04	1.309	0.292
Residual	20	428.4	21.42		
Total	22	484.5			

The critical points for $F_{2,20}$ are 3.49(5%); 5.85(1%). There is no evidence at both 1% and 5% level to reject the null hypothesis that all programmes have the same mean performance score. Therefore, we can conclude that at 1% and 5% levels, there is no evidence of significant differences in performance across the three programmes. The estimate of residual variance is 21.42 (= residual mean square). The box plot clearly demonstrates the low achievement scores in Mathematics because the number of scripts analysed is not equal to those of Biology and Physics although the difference with Biology is not significant.

Figure 3. Box Plot Achievement by programme Rater 1



Note. B

A follow-up analysis of the mean values for the three programmes is shown in table 5.

Table 5. Analysis of means by programme

Programme	Physics	Biology	Mathematics
Mean	12.78	12.33	8.8
Replicates	9	9	5

A Simple pairwise differences analysis is shown in Table 6.

Table 6. Simple Pairwise differences analysis

Comparison	**t**	**P-value**	**P**
BSc. in Physics vs BSc. in Biology	0.206	0.839	
BSc. in Physics vs BSc. in Mathematics	1.824	0.083	< 0.1
BSc. In Biology vs BSc. in Mathematics	1.618	0.121	

From the results, it can be concluded that the mean performance difference between BSc. Physics and BSc. Biology programmes is not significant. At 5% level, the difference between BSc. Physics and BSc. Mathematics is significant. So the performance in Physics is significantly better than in Mathematics. The mean performance difference between BSc. Biology and BSc. Mathematics programmes is not significant.

3.2.2 Rater 2 Scores by programme

For rater 2, the total scores for Physics were 83, Biology 113 and Mathematics 44. The total score in mathematics for both raters is the same and still much lower than Physics and Biology because only five scripts were analysed compared to the other two programmes which had nine scripts each.

The rubric scores for each programme as graded by rater 2 are shown below. It is important to note that the point of disagreement between the raters related to two scripts which show rater differences of more than 6 points and these scripts were all Physics projects. The key differences were in the scores awarded for the skills of evaluation, explanation and disposition. While rater 1 indicated 3, 3 and 3 for the three skills, rater 2 indicated 1, 1, and 2 respectively. With the rating of the two scripts by rater 3 indicated in table 5, an average of 13.7 for script 13 is attained which shows 7 points above the first rater's score and is within range with the second rater, while the average of the three scores for script 7 is 11.3 indicating 11 points above the first rater and within range with the second rater.

Table 7. Rater 2 one-way ANOVAanalysis by programme

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
PRGM	2	62.9	31.450	3.522	0.0489 *
Residuals	20	178.6	8.929		

Signif. codes:	**0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**				

The critical points for $F_{2,20}$ are 3.49(5%); 5.85(1%). There is evidence of differences when a 0.05 conventional significance level is adopted. The null hypothesis of equal mean performance scores across the three programmes is rejected. The estimate of residual variance is 8.929 (= residual mean square).
A follow-up analysis of the mean values for the three programmes is shown in Table 8.

Table 8. Analysis of means by programme

Programme	Physics	Biology	Mathematics
Mean	9.333	11.833	8.8
Replicates	9	9	5

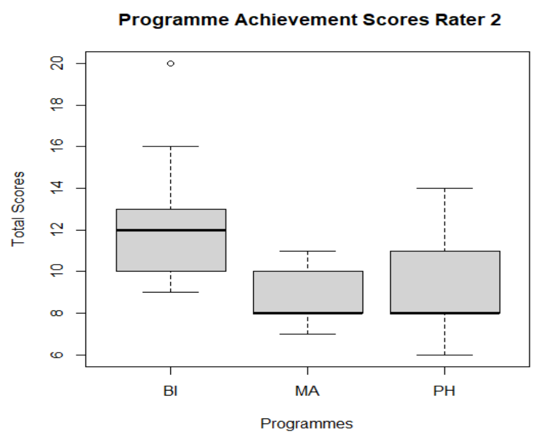
A Simple pairwise differences analysis is shown in Table 9.

Table 9. Simple Pairwise differences analysis

Comparison	t	P
BSc. In Physics vs BSc. in Biology	-1.773	< 0.1
BSc. in Physics vs BSc. in Mathematics	0.378	
BSc. In Biology vs BSc. in Mathematics	2.151	< 0.05

From this table, it can be concluded that at 10% level, the difference between BSc. Physics and BSc. Biology is significant. Therefore, the performance in Physics is not significantly better than in Biology. The performance difference between BSc. Physics and BSc. Mathematics programmes is not significant. There is a significant difference between BSc. Biology and BSc. Mathematics when a conventional level of 0.05 is adopted. Therefore, the performance in Biology is significantly better than in Mathematics.

Figure 4.Box plot



Note. Box plot representing mean scores by programme

3.3 Research question three: Is there a significant difference in performance in critical and analytic thinking skills between male and female students?

The study also sought to address the issue of gender in science and mathematics education to discover predominance of certain skills in either gender. The results show a marginal difference from the total scores.

Table 10. Analysis by gender: rater 1

Gender	Replicates (n_i)	Mean ($\bar{y}$)	Variance (s^2)	STdev (S_y)	t	P-value
Female	8	12.125	24.41	4.94	0.283	0.78
Male	15	11.53	22.27	4.72		

Table 11. Analysis by gender: rater 2

Gender	Replicates (n_i)	Mean ($\bar{y}$)	Variance (s^2)	STdev (S_y)	t	P-value
Female	8	12.25	11.36	3.37	0.659	0.517
Male	15	9.4	8.54	2.92		

Adopting a 0.01 conventional significant level, we fail to reject the null hypothesis of no difference between the overall scores of the males and females for both raters.

4. Discussion

There is overwhelming evidence from the literature and the current study that critical and analytic thinking skills are low in many senior science and mathematics students. The results show that inference, analysis and evaluation have the highest scores in that order while interpretation, explanation and disposition are the lowest in that same order.
Although this ranking is important, in the study, there is no average rater score or average overall score that reached the half point of 69 of the total 138 maximum score. The latent patterns and meanings of the evidence point to professional practice in delivering 21st century skills to learners.

In their study of Assessment of Critical Thinking Ability (ACTA) survey, White et al. (2011) show that 38.3% and 48.9% of senior science undergraduates were unable to think critically on all the three levels they tested. According to their study, the results suggest that the teaching of science does not necessarily develop critical thinking skills. This study by White et al. (2011) agrees with the current findings in terms of evidence showing that critical thinking skills are low among senior science student-teachers. This is because "thinking tends to focus on a problem's surface structure" (Willingham, 2007). Willingham (2007) further argues that teaching critical thinking skills should happen in a domain specific environment. Research projects demonstrate a student's ability to apply what has been taught and therefore much of the critical thinking "instruction" happens within subject specific teaching areas. At the research site, this 'instruction' is implicitly done in the sense that there is no course or module on critical and analytic thinking skills to which senior science students can refer. Further, even if the course or module exists, it is not called critical thinking or analytic thinking but probably falls under a broader topic and series of objectives.

Using Soper's (2015) definition of disposition, an extension of Saxton et al. (2012) definition, scripts 7, 8, 9, and 10 demonstrate close-mindedness or hostility to reason and focuses on one side of the issue and do not display an ethos of diligence. There are many instances where student authors have not done their due diligence in the current findings. For example, in the context of Saxton et al. (2012) and Soper's (2015) definitions of disposition, it becomes logical to adopt Willingham's (2007) position that acquiring critical thinking skills should be done within a specific domain – a specific teaching area in this case. When the criteria for disposition are considered, undergraduate research students can mainly acquire this skill in their subject areas to facilitate their ability to discuss and report findings. The attributes of following the evidence, adoption of a consistent point of view, awareness of alternative viewpoints (Saxton et al., 2012) and showing empathy, open-mindedness for other viewpoints, displaying an ethos of diligence, care, reasonableness, and persistence (Soper, 2015) can best be learned in the specific science domain of study which may not be learned or acquired elsewhere in order to function in a given scientific field. Although Halpern (2001) and Van Gelder (2005) disagree that teaching critical thinking is not domain specific, the value of combining these two extremes is necessary in bringing about the required teaching and learning of these skills especially in content specific areas.

The results therefore imply and suggest that senior science students need to be taught metacognition – being aware of how we think (Thomas, 2011) in an explicit and deliberately planned manner. This is what others term 'thinking about thinking' (Murawski, 2014). Walsh and Hardy (1999) showed that based on the major of study, there are differences in disposition toward critical thinking among college students. Thomas (2011) suggests how critical thinking can be developed and assessed in the first year of a student's academic life through the following:

a. Method 1 comprises evaluating the quality of sources on the internet
b. Method 2 comprises analysis and synthesis of an argument
c. Method 3 developing critical thinking and logical thinking using the immediate feedback assessment technique (Epstein Education, 2010)
d. Method 4 using rubrics and reflection as a means of self-regulation.

The three levels of the Assessment of Critical Thinking Ability (ACTA) surveyed and reported by White et al. (2011) could be adopted early on in teacher training of science students to prepare them for effective critical thinking in academic writing and argumentation. These three ability levels are:

a. Integrating conflicting studies into a unified conclusion
b. Designing experiments to resolve ambiguities in particular studies
c. Conjecturing other interpretations of particular studies

Dunbar (2006) has demonstrated that rubrics can identify changes and improvements in course delivery and design. In terms of skills analysis, the study shows that inference had the highest score and disposition in critical thinking scored far less than the other five skills. The reasons for this are unclear but the existing evidence from other studies indicates that disposition varies with gender. According to Shubina and Kulakli (2019) and Walsh and Hardy (1999), females have higher level disposition strengths in the areas of open-mindedness and cognitive maturity.

The practical implications of the current findings suggest a need to integrate the 21st century skill of critical and analytic thinking in research methodology and writing tasks in content or subject areas and explicitly teach these skills. The practical suggestions by Thomas (2011) and White (2011) above could be used as an entry point for the mathematics and science education programmes at the research site.

5. Conclusion

This study has demonstrated the problem of the need for evidence-based findings, the need for deliberate instruction in the *discussion of findings* and the potential value that critical and analytic thinking adds to transforming 21st century education. The data shows that there is evidence that mathematics students' performance is lower than that of biology and physics students. Further the data indicates that there is no significant difference in performance scores of male and female students.

6. Acknowledgements

We wish to thank Professor Overson Shumba for initiating the direction of the research and continued intellectual support. He has championed the integration of education for sustainable development and particularly action research and spearheading the 21st century skills and competences agenda for science and mathematics education. His insights, comments, and suggestions have significantly motivated the paper.

References

Abrami, P.C., Bernard, R.M., Borokhovski, E., Wade, A., Surkes, M.A., Tamim, R., & Zhang, D. (2008). Instructional Interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis. Review *of Educational Research*, 78(4), 1102-1134. https://doi.org/10.3102/0034654308326084

Anderson, L. & Krathwohl, D. (Eds.) (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives.* New York: Longman.

Bagheri, F. & Ghanizadah, A. (2016). Critical thinking and gender differences in academic self-regulation in higher education. *Journal of Applied Linguistics and Language Research*, 3(3), 133-145.

Behar-Horenstein, L.S & Niu, L. (2011). Teaching Critical Thinking Skills in Higher Education: A Review of the Literature. *Journal of College Teaching & Learning*, 8(2), 25 – 41.

Campbell, A. (2005). Application of ICT and rubrics to the assessment process where professional judgment is involved: The features of an e-marking tool. *Assessment & Evaluation in Higher Education*, 30 (5), 529-537.

Changwong, K., Sukkamart, A., & Sisan, B. (2018). Critical thinking skill development: Analysis of anew learning management model for Thai high schools. *Journal of International Studies*, 11(2), 37-48. https://doi.org/10.14254/2071-8330.2018/11-2/3

Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education*. (6th ed.). London: Routledge – Taylor & Francis Group.

Docherty, M. & Smith, R. (1999). The case for structuring the discussion of scientific papers: Much the same as that for structuring abstracts. *British Medical Journal*, 318 (7193), 1224-1225. http://doi.org/10.1136/bmj.318.7193.1224

Dunbar, N.E., Brooks, C.F. & Kubicka-Miller. (2006). Oral communications skills in higher education: Using a performance-based evaluation rubric to assess communication skills. *Innovative Higher Education*, 31(2), 115-128.

Facione, P.A. (1990a). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. Millbrae, CA: California Academic Press.

Facione, P.A. (1990b). *California Critical Thinking Skills Test (CCTST) Forms A & B.* Millbrae, CA: California Academic Press.

Facione, P. A (2000). The disposition toward critical thinking: Its character, measurement, and relationship to critical thinking skill. *Informal Logic*, 2(1), 61-84. https://doi.org/10.22329/il.v20i1.2254

Facione, P.A. (2010). Critical thinking: What it is and why it counts. Insight Assessment. [Online] Available: http://www.insightassessment.com/pdf_files/what&why2006.pdf

Foote, M. (2009). The proof of the pudding: How to report results and write a good discussion. *Chest*, 135(3), 866-868. https://doi.org/10.1378/chest.08-2613

Halpern, D.F. (2001). Assessing the effectiveness of critical thinking instruction. *The Journal of General Education*, 50(4), 270-286.

Kennedy, M., Fisher, M. B., & Ennis, R. H. (1991). Critical thinking: Literature review and needed research. In L. Idol & B.F. Jones (Eds.), Educational values and cognitive instruction: Implications for reform (pp. 11-40). Hillsdale, New Jersey: Lawrence Erlbaum & Associates.

Knight, L.A. (2006). Using rubrics to Assess Information Literacy. *Reference Services Review*, 34(1), 43-55.

Lai, E.R. (2011). *Critical Thinking: A Literature Review Research Report*. Pearson. http://www.pearsonassessments.com/research

McCombes, S. (2019). How to write a discussion section. https://www.scribbr.com/dissertation/discussion/

Ministry of Education. (2013). *Zambia Education Curriculum Framework*. Government Printers.

Murawski, L. M. (2014) Critical Thinking in the Classroom…and Beyond. *Journal of Learning in Higher Education*, 10(1), 25-30.

Partnership for 21st Century Skills. (2009). *P21 Framework Definitions. Partnership for 21st Century Skills*. http://www.21stcenturyskills.org

Saxton, E., Belanger, S., & Becker, W. (2012). The Critical Thinking Analytic Rubric (CTAR): Investigating Intra-Rater and Inter-Rater Reliability of a Scoring Mechanism for Critical Thinking Performance Assessments. *Assessing Writing*, 17(4) 251-270. https://www.academia.edu/15516218/The_Critical_Thinking_Analytic_Rubric_CTAR_Investigating_intra-rater_and_inter-rater_reliability_of_a_scoring_mechanism_for_critical_thinking_performance_assessments

Thomas, T. (2011). Developing First Year Students' Critical Thinking Skills. *Asian Social Science*, 7(4), 26-35.

Van Gelder, T. (2005). Teaching critical thinking: Some lessons from cognitive science. *College Teaching*, 53(1), 41-48.

Vieira, R.F, de Lima, R.C. & Mizubuti, E.S.G. (2019). How to write the discussion section of a scientific article. *Acta Sci*., Agron, 41. https://doi.org/10.4025/actasciagron.v41i1.42621

Wallwork, A. (2016). *English for writing research papers*. New York, NY: Springer.

Watson, G., & E.M. Glaser. (1952). *Watson-Glaser Critical Thinking Appraisal.* New York: Pearson Education.

White, B., Stains, M., Escriu-Sune, M., Medaglia, E., Rostamnjad, L., Chinn, C., & Sevian, H. (2011). A Novel Instrument for Assessing Students' Critical Thinking Abilities. *Journal of College Science Teaching*, 40 (5), 102-107.

Walsh, C. & Hardy, R. (1999). Dispositional differences in critical thinking related to gender and academic major. *Journal of Nursing Education*, 38 (4), 149 – 155.

Zetriuslita, H.J., Ariawan, R., & Nufus, H. (2016). Students' Critical Thinking Ability: Description based on Academic Level and Gender. *Journal of Education and Practice*, 7 (12), 154-164.