2020

Mathematics Cognitive and Content Abilities Across The Vincentian Student Population

Brendalee Rorena Cato
Walden University

Walden University

College of Education

This is to certify that the doctoral dissertation by

Brendalee R. Cato

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Bonnie Mullinix, Committee Chairperson, Education Faculty
Dr. John Flohr, Committee Member, Education Faculty
Dr. Kimberley Alkins, University Reviewer, Education Faculty

Chief Academic Officer and Provost
Sue Subocz, Ph.D.

Walden University
2020

Abstract

Mathematics Cognitive and Content Abilities Across The Vincentian Student Population

by

Brendalee R. Cato

MPhil, University of the West Indies, 1998

BEd, University of the West Indies, 1996

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Assessment Evaluation and Accountability

Walden University

November 2020

Abstract

Mathematics achievement is a key component of student overall academic achievement. However, many students from Saint Vincent and the Grenadines (Vincentian students) continue to perform poorly on the regional Caribbean Secondary Examination Certificate (CSEC) mathematics examination. This poor mathematics performance is a concern for education stakeholders. The purpose of this quantitative, nonexperimental study was to explore the extent to which the CSEC mathematics scores of high-scoring Vincentian students versus low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning differ across the content domains of algebra; geometry; measurement; statistics; and relations, functions, and graphs (RFG). The theoretical foundation for the study was Bloom's taxonomy of educational objectives. The study used a cross-sectional design and archival data. The sample was composed of 370 students. Two-way multivariate analysis of variance (MANOVA) and follow-up 2-way analysis of variance were computed to provide answers to the research question. Based on the MANOVA, there was a statistically significant interaction effect between levels of knowledge and levels of reasoning for measurement scores. Additionally, there were significant main effects for each cognitive domain and algebra, geometry, measurement, and RFG. The findings of the study contribute to positive social change by providing teachers, administrators, and education policy makers in Saint Vincent and the Grenadines with insights into the influence of cognitive abilities on student mathematics achievement so that they could identify students who may be at risk for learning difficulties in mathematics and better plan intervention strategies for remediation.

Mathematics Cognitive and Content Abilities Across The Vincentian Student Population

by

Brendalee R. Cato

MPhil, University of the West Indies, 1998

BEd, University of the West Indies, 1996

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Assessment Evaluation and Accountability

Walden University

November 2020

Dedication

I would like to dedicate this study to the students of the Caribbean, and particularly students from Saint Vincent and the Grenadines who struggle with learning mathematics and passing mathematics examinations. Like you, I too struggled with mathematics in my early academic journey, but my love for the subject, and my early recognition of its importance to daily existence and a successful career, motivated me to persevere. If I can do it, you can too. Many of you may feel like you are a failure because you did not achieve a certificate in mathematics, but in many cases, you did not fail, the system failed you. Some of the teaching strategies used by many teachers of your teachers encourage rote learning, they do not cater to individual learning needs, learning styles and interests, especially of digital novices like yourselves; neither do they foster the development of critical thinking and problem-solving skills, which are essential to success in mathematics. Hence, your teachers may have failed to provide you with the necessary tools and strategies for mathematics achievement. Through this research, I have provided recommendations to assist teachers in improving mathematics pedagogy and making learning meaningful and enjoyable to students.

Acknowledgments

Firstly, I would like to thank the Almighty God for his blessings of knowledge and strength to successfully complete this noble journey. I couldn’t do it without your grace and mercy. Thank you, Lord!

I would also like to thank my children; Kevin, Khea, and Keana for their patience and understanding over the years as my time was consumed pursuing my dream. Keana was always curious to find out what was keeping me so busy, the long nights and lockdown weekends even after a hard work week. I know that you are as happy as I am that this part of the journey has concluded, and as promised, we will be able to do more fun things together.

To my great team of committee members, Dr. Bonnie Mullinix, Dr. John Flohr, and Dr. Kimberley Alkins, I say a hearty thank you. My Chair, Dr. Bonnie Mullinix, is a phenomenal mentor. Her amazing qualities have contributed in a great way to my success. She is a patient listener, and very supportive and respectful of others’ opinion. Throughout the dissertation process, she has provided clear guidance, unflinching support, and timely, detailed feedback, and has always expressed confidence in my ability. Dr. Flohr has provided great leadership, especially in the methodological aspect of the study. His feedback was timely and meticulous. His input has contributed to the success of my study. Dr. Alkin’s expertise in the field is commendable. She has provided comprehensive and timely feedback that have contributed to high quality of my study. I was blessed to have had a knowledgeable and capable team to assist in navigating my journey and making it a pleasant one. To my committee, thank you.

My mom, Eugenia Edwards, my dad, Samuel Edwards, and my husband, Jeffrey Cato, neither of you made it to see my accomplishment. To my siblings, Glenroy, Brensley, Andrea and Susanna, I share my success with you. Thanks to Ms Phillisha Chance who stepped in to assist with other duties so I could dedicate my time to concentrated writing. My friends Mr. Al Barrack, Dr. GraceAnne Jackman, Ms. Cyndra Ramsundar, and Ms. Patricia Clarke, thank you for your support, encouragement, and motivation, as well as the level of confidence you displayed in my ability to succeed. I couldn’t let you down.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction to the Study

Introduction

In this study, I investigated whether students from Saint Vincent and the Grenadines (Vincentian students) classified as high scoring versus those classified as low scoring in the cognitive domain of knowledge, comprehension, and reasoning performed differently in the content domains of algebra; geometry; measurement; and relations, functions, and graphs (RFG) in the 2017 May/June Caribbean Secondary Education Certificate (CSEC) mathematics examination. Through this research, I endeavored to fill a gap in the literature relating to the lack of research on cognitive abilities and mathematics achievement of students in the Caribbean, in general, and in Saint Vincent and the Grenadines in particular. I hoped to provide an in-depth understanding of the influence of the cognitive abilities of knowledge, comprehension, and reasoning on Vincentian student mathematics achievement in the content domains of algebra, geometry, measurement, statistics, and RFG. This insight may help promote positive social change by influencing policy decisions regarding mathematics education in Saint Vincent and the Grenadines. Teachers, administrators, and education policy makers could use the insights gained to identify students who are at risk for learning mathematics and plan intervention strategies for remediation, with the view to supporting mathematics pedagogy. The expected improvement in student learning should increase their career options and ultimately lead to a better quality of life.

In this chapter, I discuss the background of the study, including a brief summary of the research literature relating to the topic and a description of the gap in knowledge that the study addresses and why the study is needed. I describe the background to the study, including the developmental history and structure of the CSEC mathematics examination. I also include a statement of the research problem and its relevance and significance to the discipline, as well as the identification of a meaningful gap in current research literature that I sought to address in the study. Following the information on the gap, is a statement of the purpose of the study and the nature and type of the study, as well as the independent and dependent variables. I then present the research question and hypothesis, followed by a description of the theoretical framework of the study and an explanation of how the framework relates to the study approach and research questions. I give an outline of the nature of the study; this includes a rationale for the design selected, and operational definitions for the variables used in the study. I then present the assumptions of the study, including their likely effect on the meaningfulness of the study. Following the assumptions section, I provide a description of the scope and delimitations of the study; including treatment of the research questions, issues of internal validity, and the boundaries of the study that could affect the external validity of the study. The limitations of the study pertaining to methodological weaknesses, possible biases that could influence the study outcomes, and how these will be addressed follow. Finally, I discuss the significance of the study, including how the study will advance knowledge in the discipline, influence policy and practice, and have potential implications for social change. The chapter concludes with a summary of the main points and a transition to Chapter 2.

Background

Mathematics is a multifaceted system of complex relationships that involves and invokes reasoning (Morsanyi, Prado, & Richard, 2018). According to Soni and Kumari (2017), mathematics is a skill that is indispensable in all facets of life. Mathematics achievement is a major component of overall academic achievement (Vista, 2016). Mathematics plays a pivotal role in nation building and a vital tool for understanding and predicting future phenomenon (Bassey, Joshua, & Asim, 2009). Bassey et al. (2009) summarized the importance of mathematics education as “mathematics education is to a nation what protein is to a young organism” (p. 56). Mathematics is considered an essential 21st-century competency for leading a fulfilling life and functioning effectively in a dynamic society that is becoming progressively “quantified” (Cragg, Richardson, Hubber, Keeble, & Gilmore, 2017; Karakolidis, Pitsia, & Emvalotis, 2016).

A high level of mathematics proficiency is critical for success at the individual level as well as societal level (Lipnevich, Preckel, & Krumm, 2016). At the societal level, mathematics is considered to be fundamental to the advancement of economic development, particularly in developing countries (Bosman & Schulze, 2018). Moses and Cobb (2001) shared that mathematics and science literacy is crucial in liberating and stabilizing society and affording people economic access and full citizenship. The authors believed that mathematics literacy and economic access will give hope to the young generation and that they will close the knowledge gap and prepare citizens for the future. Competence in mathematics is critical to the workforce in science, technology, engineering, and mathematics (STEM) disciplines and to international leadership (Jordan, Glutting, & Ramineni, 2010). In a competitive global economy, a workforce that is competent in STEM is likely to guarantee future economic prosperity (Panizzon et al., 2018). A lack of proficient persons in mathematics-related disciplines will result in economic disadvantages (Lipnevich et al., 2016).

On an individual level, success in mathematics is related to health, well-being, satisfaction with life, longevity, employability, and wages (Lipnevich et al., 2016; Reyna & Brainerd, 2007). Basic knowledge of high school mathematics is required for entry-level employment in both private and public sectors, as well as the army (Erden & Akgul, 2010). Mathematics creates greater career options, particularly in high-paying fields such as engineering, information technology, and finance (Mji & Makgato, 2006). Moreover, mathematics proficiency is essential for performing task of everyday living, including decision making (Cragg et al., 2017; Reyna & Brainerd, 2007). Achievement in mathematics is inextricably linked to future career opportunities. In contemporary societies, achievement in mathematics can be a gateway to personal and economic success (Primi, Bacherini, Beccari, & Donati, 2020; Waxman, 2020).

Given the significant role that mathematics plays in student overall academic achievement (O'Connell, 2018), education policy makers in the Caribbean made the subject compulsory for students taking CSEC examinations at the secondary level. The government of Saint Vincent and the Grenadines, in its support for a better education system, has invested a substantial portion of its budget to the education sector (Prince, 2018). Despite these efforts, Vincentian students continue to perform poorly in the CSEC mathematics examination. This poor performance is a major concern for education stakeholders in Saint Vincent and the Grenadines. Although a minuscule number of Vincentian students perform well on the CSEC mathematics examination, most students continue to fail the examination. Analysis of the annual CSEC mathematics examination results for 10 years, 2008 to 2017, shows that students consistently scored below 50% of the available marks on the examination. Performance is generally poor in all content domains, and students score lowest in the cognitive domain of reasoning (Caribbean Examinations Council [CXC], 2018).

Notwithstanding the persistent poor performance of students on the CSEC mathematics examination, the problem has not been formally investigated; hence, there seems to be an apparent lack of knowledge among Caribbean educators regarding possible factors that contribute to such poor performance. In this study, I sought to fill a gap in knowledge regarding the influence of Vincentian students' cognitive abilities on their achievement in the CSEC mathematics examination. Such insights may help educators to better understand possible reasons for such poor performance and identify those students who may be at risk for poor mathematics achievement and plan intervention strategies for remediation.

Regional Mathematics Assessment

Regional assessments in the Caribbean are developed and administered by the CXC. The main role of the CXC is to provide assessments and certifications of Caribbean students mainly at the secondary level. However, the CXC also develops contracted examinations for transition of students from primary to secondary level, as well as professional licensure examinations. The CXC was established in 1973 by a group of educators from universities and colleges in the Caribbean with a view to transforming education in the Caribbean (Bryan, 2014). Education policy makers believed that the General Certificate of Examination (GCE) offered by Cambridge, England, did not meet the needs of Caribbean students as the syllabi represented the culture of Britain and not of the Caribbean. Also, the syllabi were not geared toward the economic and social development of the Caribbean (Bryan, 2014). Regional educators believed that establishing a Caribbean-focused examination would be an asset to the Caribbean. During the 1960s and 1970s, a number of Caribbean governments substantially increased educational opportunities at the primary, secondary, and tertiary levels through sizable increases in their education budgets. Some countries introduced free secondary education, which resulted in an increased demand for secondary school education.

The Caribbean political directorate endorsed the establishment of a CXC. In 1964, a working group comprising ministers of education of Barbados, Jamaica, Guyana, and Trinidad and Tobago met in Barbados to discuss the composition and functions of the proposed CXC. After several meetings and discussions, the final agreement establishing the Council was reached in 1972. A committee comprising Caribbean educators embarked on the “Caribbeanisation” of syllabi and examinations. The committee employed experts from Cambridge to guide the process. They agreed to locate the headquarters of the CXC in Barbados and a Western Zone office in Jamaica. The first suite of examinations, including mathematics, was offered to 13 Caribbean countries in 1979 (Bryan, 2014). Currently, the CXC offers 30 subjects at the CSEC level, and 33 subjects at the Caribbean advanced proficiency level (CAPE) to 19 Caribbean countries.

The 19 countries are Antigua and Barbuda, Anguilla, Barbados, Belize, British Virgin Islands, Cayman Islands, Dominica, Grenada, Guyana, Jamaica, Montserrat, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Trinidad and Tobago, Turks and Caicos, Suriname, Saba, and Saint Maarten (CXC Annual Report, 2017).

Structure of the CSEC Mathematics Examination

The CSEC mathematics examination comprises two components: Paper 01, a 60-item compulsory multiple-choice paper that is worth 60 marks (points); and Paper 02, a constructed response paper that is worth 120 marks. Paper 02 is divided into two sections. Section 1 comprises eight compulsory questions and Section 2 comprises three optional questions worth 15 marks each. Students are required to answer two of the three optional questions. Both papers assess competencies in 10 content domains across three cognitive domains. The content domains are computation, number theory, sets, consumer arithmetic, measurement, algebra, statistics, relations, functions and graphs, geometry and trigonometry, and vectors and matrices. The cognitive domains are knowledge, comprehension and reasoning. Table 1 shows the structure of the examination by cognitive domain and Table 2 shows the structure of the examination by content domain.

Table 1

Structure of CSEC Mathematics Examination by Cognitive Domain

Cognitive domain	Paper 01 No. of Marks	Paper 02 No. of Marks	Total No. of Marks
Knowledge	18	36	54
Comprehension	24	48	72
Reasoning	18	36	54
Total	60	120	180

Note. Adapted from “CSEC, 2018,” p. 5

Table 2

Structure of CSEC Mathematics Examination by Content Domain

Content domain	Paper 01 Multiple-choice	Paper 02 Constructed response
Compulsory	No. of marks	No. of marks
Number theory	4	10
Computation	6	-
Consumer arithmetic	8	10
Sets	4	5
Measurement	8	10
Statistics	6	15
Algebra	9	10
Relations, functions and graphs	5	10
Geometry and trigonometry	9	20
Optional		
Algebra, relations, functions, and graphs	-	15
Geometry and trigonometry	-	15
Vectors and matrices	-	15

Note. Adapted from "CSEC," 2008, pp. 2-3.

The ratio of the weighting of Papers 01 to 02 is 1:2. Paper 01, the multiple-choice paper, contributes one-third to the overall weighting of the subject, whereas Paper 02, the constructed response paper, contributes two-thirds to the overall weighting.

Problem Statement

The poor performance of students in mathematics is a global problem (di Gropello, 2017). According to the National Center for Education Statistics (NCES, 2016), 60% of American fourth- to 12th-grade students performed below the proficiency level in the National Assessment of Education Progress mathematics examination. Poor mathematics achievement is a recognized problem for Caribbean educators and policy makers who continue to lament the poor mathematics performance of Caribbean students and the need to confront and arrest the problem (Bruns & Luque, 2015; Cumberbatch, 2016; Jules, 2012; Leacock, 2015; Monteith, 2016; Quinn-Leandro, 2011; Quinn-Leandro, 2012; Reid, 2011; Sodha, 2012). Data for the 10-year period, 2008 to 2017, show that more than 60% of Caribbean students fail the CSEC mathematics examination every year. Generally, Caribbean students achieve very low scores in all content domains, but particularly in algebra, geometry, measurement, statistics, and RFG (CXC, 2018). The scores on the content domains ranged 19% to 48% during the 5-year period, 2013 to 2017. In addition to the overall poor performance in mathematics, students' generally score lowest on the reasoning profile, demonstrating their inability to engage in higher-order thinking skills (CXC, 2018). Although the performance of Caribbean students is generally poor, the performance of Vincentian students is alarming. The mean percentage score of Vincentian students is lower than the mean percentage scores for the Caribbean.

Overall, Saint Vincent and the Grenadines consistently ranks in the lowest three of the 19 Caribbean countries in CSEC mathematics examination (CXC, 2018).

Mathematics education comprises two dimensions: content domains, or tasks and cognitive domain or skills required to solve the tasks (Männamaa, Kikas, Peets, & Palu, 2012). A search of the literature reveals that most research on cognitive abilities and mathematics achievement have been conducted at the primary level (Geary, 2011; Primi, Ferrão, & Almeida, 2010; Wong & Ho, 2017). There is a lack of research on cognitive abilities, as defined by based on Bloom, Engelhart, Furst, Hill, and Krathwohl (1956) taxonomy, and mathematics achievement at the secondary level, particularly in the Caribbean. Most of the research on cognitive abilities and mathematics achievement have been conducted in the United States and are based on the Cattel-Horn-Carroll (CHC) theory of human cognitive abilities, which focuses on broad cognitive abilities and general intelligence. Although Trends in Mathematics and Science Study (TIMSS) uses Bloom taxonomy (Bloom et al., 1956) in its international assessment of Grade 4 and Grade 8 students, the assessment does not include students from Saint Vincent and the Grenadines. Hence, an in-depth understanding of the influence of students' cognitive abilities of knowledge, comprehension, and reasoning on their mathematics achievement in the content areas of algebra, geometry, measurement, statistics, and RFG as assessed in the CSEC mathematics examination is needed. This knowledge may help teachers, administrators, and education policy makers in Saint Vincent and the Grenadines identify students who may be at risk for poor mathematics achievement so that they could better target instructional areas for remediation that will support mathematics pedagogy among Vincentian students.

Purpose of the Study

The purpose of this quantitative, nonexperimental study is to determine the extent to which the CSEC mathematics scores of high scoring Vincentian students versus low scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning differ across the content domains of algebra, geometry, measurement, statistics, and RFG. The theoretical framework was Bloom's taxonomy of educational objectives (Bloom et al., 1956). Bloom et al. (1956) taxonomy comprises six levels of cognitive skills: knowledge, comprehension, application, analysis, synthesis, and evaluation (Granello, 2001). The CSEC mathematics examination is designed based on Bloom's taxonomy. The first two cognitive domains in the mathematics examination mirror the first two levels of Bloom et al. (1956) taxonomy, whereas the third domain, reasoning, in the CSEC mathematics examination encapsulates the other levels of the taxonomy from application to evaluation.

Variables

There were four independent variables, each with two levels, and five dependent variables.

Independent variables. The independent variables included the cognitive domain (knowledge, comprehension, and reasoning) and the high and low categories of performance (CoP) groups.

Dependent variables. The dependent variables were the scores in the content domain for algebra, geometry, measurement, statistics, and RFG.

Research Question and Hypotheses

Research Question: How do the CSEC mathematics scores of high-scoring Vincentian students versus low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning, differ across the content domains of algebra, geometry, measurement, statistics, and RFG?

H_0: There are no differences in the CSEC mathematics scores between high-scoring Vincentian students and low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning, across the content domains of algebra, geometry, measurement, statistics, and RFG.

Ha: There are differences in the CSEC mathematics scores between high-scoring Vincentian students and low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning, across the content domains of algebra, geometry, measurement, statistics, and RFG.

Descriptive statistics, correlation, two-way multivariate analysis of variance (MANOVA) and follow-up two-way analysis of variance (ANOVA) were used to determine whether students classified as high scoring and low scoring in the cognitive domains (knowledge, comprehension, and reasoning) would performance differently in CSEC mathematics examination, measured by the scores on five content domains (areas: algebra, geometry, measurement, statistics, and RFG).

Theoretical Framework

The theoretical base for this study was Bloom's taxonomy of educational objectives (Bloom et al., 1956). Bloom's taxonomy is a hierarchical organization of six global educational objectives (Ursani, Memon, & Chowdhry, 2014). The objectives, defined in behavioral terms, are knowledge, comprehension, application, analysis, synthesis, and evaluation. The hierarchy represents mental processes from simple to complex, concrete to abstract, and mastery of each simple category is a prerequisite to mastering the next complex category (Anderson, Krathwohl, & Airasian, 2001; Lipscomb, 1985). Bloom's taxonomy is a framework developed to provide instructors with a systematic assessment of student behavior as a result of participating in an educational experience. The taxonomy was intended to form a universal language among teachers and assist them in creating testing materials that more accurately assess their curriculum aim (Bertucio, 2017). Bloom's taxonomy is a taxonomy of general competence that serves to guide educational objectives and has been used to improve pedagogy and assessment methods in many disciplines (Ursani et al., 2014).

Bloom et al.'s (1956) taxonomy of educational objectives has provided a foundation for the understanding of learning outcomes and a platform for the development of other taxonomies, including the Revised Bloom's Taxonomy (Anderson et al., 2001); Marzano's New Taxonomy (Marzano & Kendall, 2006); and Mathematical Wellbeing (Clarkson, Bishop, & Seahs, as cited in Irvine, 2017). Despite the development of more recent taxonomies, the original Bloom's taxonomy (Bloom et al., 1956) has been used extensively by educators to identify and delineate tasks involving higher-order and lower-order thinking skills (Irvine, 2017). Bloom's taxonomy has also been found to be exceptionally helpful in providing clarity in designing the teaching process by structuring and sequencing educational objectives in needs assessment, lesson planning, and assessment (Ramirez, 2017).

In addition to classroom practitioners, many large-scale assessments have been modelled based on Bloom et al.'s (1956) taxonomy of educational objectives. TIMSS uses three levels of cognitive domains (knowing, applying, and reasoning), to develop a content-by-process matrix to create mathematics assessment in the content domains of number, algebra, geometry, and data and probability, for eighth-grade students internationally. The description of TIMSS's cognitive domains closely match the first three levels of Bloom's taxonomy of educational objectives. The CXC has adopted and used Bloom's taxonomy in the creation of its regional 11th-grade, CSEC mathematics examination. The CSEC mathematics assessment follows a similar content-by-process matrix as the TIMSS's eight mathematics assessment as outlined in the TIMSS assessment framework (Mullis & Martin, 2019). The CSEC assessment includes the content domains of computation, number theory, consumer arithmetic, sets, measurement, statistics, algebra, relations, functions and graphs, and geometry and trigonometry and the cognitive domains of knowledge, comprehension, and reasoning.

In this study, the assessment matrix comprised the three cognitive domains (knowledge, comprehension, reasoning) and five content domains (algebra, geometry, measurement, statistics, and RFG). These content domains represent the areas of poorest performance for Caribbean students including Vincentian students, on the 2017 May/June CSEC mathematics examination.

Nature of the Study

In this quantitative study, I used a cross-sectional design and archival (secondary) data. The source of the data was the CXC database. The data were comprised of Vincentian students' scores in the 2017 May/June CSEC mathematics examination. A cross-sectional design allowed for data to be collect at one point in time. In addition to reducing time and cost, I was able to collect a larger sample than would be feasible with other research designs. Using a cross-sectional design also allowed me to investigate one Caribbean country and generalize the findings to other Caribbean countries with similar characteristics. There were four independent variables that are categorical variables: three cognitive domain (knowledge, comprehension, and reasoning) variables and the CoP variable. Each independent variable had two levels: high-scoring students and low-scoring students. There were five dependent variables that were measured at the continuous levels: algebra scores, geometry scores, measurement scores, statistics scores, and RFG scores. I used SPSS version 25 to analyze the data. The data analysis included descriptive statistics, correlation, two-way MANOVA statistical analysis for differences between groups, and two-way ANOVA. I used the two-way MANOVA and two-way ANOVA to test the null hypothesis (*H*o)—there are no differences in the CSEC mathematics scores between high-scoring Vincentian students versus low-scoring Vincentian students in the three cognitive domains of knowledge, comprehension, and reasoning, across the five content domains of algebra, geometry, measurement, statistics, and RFG.

Definition of Terms

Algebra: A way of thinking that involves the analysis of mathematical situations and generalization of models devised from the application of concepts and skills (National Council of Teachers of Mathematics [NCTM], 2006). Introductory algebra involves the recognition of patterns and the use of symbols and expands to involve the use of numbers (Lee, Collins, & Melton, 2016).

Category of performance: Students classified as high scoring and low scoring based on their score in the combined five mathematics content domains in the 2017 May/June CSEC mathematics examination (algebra, geometry, measurement, statistics, and RFG).

Cognitive domain: Classification of questions or tasks based on the kind of cognitive demand (CSEC, 2008). Cognitive domain is also referred to as cognitive ability.

Comprehension: Algorithmic thinking involving translation from one mathematical mode to another. The application of algorithms to familiar problem situations (CSEC, 2008).

Geometry: The study of properties, relationships, and transformations of spatial objects within an interconnected network of concepts and representational systems (Crompton, Grant, & Shraim, 2018).

High-scoring students: Students with scores from 48 to 94 on the combined content domains in the 2017 May/June CSEC mathematics examination. This group represented students who scored above 50% on the exam.

Knowledge: The recall of rules, procedures, definitions, and facts (CSEC, 2008).

Low-scoring students: Students with scores from 0 to 47 on the combined content domains in the 2017 May/June CSEC mathematics examination. This group included students who scored 50% or below on the exam.

Mathematics content domain: Strands, area, or concepts in mathematics. *Content domain* is also referred to as *content area* and *content strand* (CSEC, 2008).

Measurement: A foundation concept in mathematics that is required for day-to-day functioning in the world (Hurrell, 2015).

Reasoning: Involves the translation of nonroutine problems into mathematical symbols and then choosing suitable algorithms to solve the problems. Reasoning also involves combining algorithms to solve problems and using algorithms in reverse order, and making inferences and generalizations, justifying results and statements, and analyzing and synthesizing (CSEC, 2008).

Relations, functions, and graphs (*RFG*): An area in mathematics associated with collecting and interpreting numerical information and communicating important relationships (Larson & Whitin, 2010).

Statistics: A branch of mathematics that deals with the collection, analysis, interpretation, and presentation of masses of numerical data (Capaldi, 2019).

Assumptions

There were five assumptions associated with the study. The first assumption was that the 2017 May/June CSEC mathematics examination has content-related evidence of validity. That is, the assessment tasks adequately represent the content measured as defined in the test blueprint or table of specifications. The second assumption was that the 2017 May/June CSEC mathematics examination has construct-related evidence of validity. That is, there is empirical evidence that the inferred constructs exist and are accurately measured in the test (Popham, 2002). The third assumption was that the cognitive domains of knowledge, comprehension, and reasoning are accurately and consistently operationalized in each test item, based on their definition in the CSEC mathematics syllabus. The fourth assumption was that the test items accurately reflect the identified content domains of algebra, geometry, measurement, statistics, and RFG, as outlined in the CSEC mathematics syllabus. The fifth assumption was that students' tests were accurately scored and the scores were accurately reported. That is, teachers consistently applied the scoring rubric in the scoring of students' work, and the scores were accurately recorded and reported.

Scope and Delimitations

The problem that I addressed in this study was the poor performance of Vincentian students in the CSEC mathematics examination. In addressing this problem, I explored the extent to which the CSEC mathematics scores of high-scoring Vincentian students versus low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning differ across the content domains of algebra, geometry, measurement, statistics, and RFG. The study was quantitative in nature and I used a nonexperimental, cross-sectional design. I explored the research problem using secondary data. I sought to provide educators and policy makers in Saint Vincent and the Grenadines with insights into the role of cognitive abilities in the learning of mathematics with a view to positively influencing mathematics pedagogy, and ultimately lead to improved student performance in the CSEC examination. However, there are certain scope and delimitations to the study, including threats to internal validity and external validity.

According to Jackson, O'Callaghan, and Adserias (2014), threats to the internal validity of a study are issues or problems with procedures or participants that can compromise the inferences that are drawn from the study. Threats to internal validity in a cross-sectional design include measurement errors that could result in spurious findings, or common-method variance (CMV) and erroneous casual inference (Jackson et al., 2014). Jackson et al. (2014) referred to CMV as variance attributable to the method used to measure the construct, rather than to the construct being measured. CMV may inflate or deflate the correlation among research variables, thereby threatening the validity of the conclusions drawn about the relationships between the measures of different constructs (Reio, 2010). These measurement methods may include using a single rater, item characteristics, item context, and measurement context (Campbell & Fiske, 1959; Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). As a potential source of measurement error, CMV in quantitative studies can be controlled by strengthening the procedural design of the study, and by using statistical controls (Podsakoff et al., 2003; Rindfleisch, Malter, Ganesan, & Moorman, 2008). In this study, I minimized measurement errors by using the same measurement instrument, in the form of an examination. All the students in the sample wrote the same mathematics examination, and the teacher used the same scoring rubric to mark the examination under the same conditions. To ensure reliability in marking, two markers marked each sample of script. Also, the data manager used the same method to retrieve all the students' scores from the CXC's database. To strengthen the procedural design of the study, I requested that the data manager perform data cleaning by removing all indefinable student information prior to delivering the data to me. Prior to data analysis, I tested the assumptions of the statistical tests to guide my interpretation and reporting of the data. I also performed follow-up statistical tests to ensure that the observed differences were accurately identified. To reduce selection bias and chance bias as sources of internal validity, I applied the G*statistics to determine an adequate sample size (n = 40) for the study. However, I used a larger sample size (n = 370) to ensure that the sample requirements were met for a small effect size and control for both the Type 1 error probability α and the Type 2 error probability $1-\beta$ (Mayr, Erdfelder, Buchner, & Faul, 2007).

Threats to external validity are problems that threaten the generalizability of the findings of one study to other setting, persons, and situations (Frankfort-Nachmias, Nachmias, & DeWaard, 2015). To reduce threats to external validity, I used stratified random sampling to ensure that subgroups of high-scoring students and low-scoring students in the sample of Vincentian students represent the subgroups of high-scoring students and low-scoring students in the Vincentian student population. I have generalized the findings of the study to the total sample studied and not to any subgroups. Additionally, I generalized the findings of the study only to the cohort of students who wrote the examination in May/June 2017 and no cohort who wrote the examination in any other sitting.

A delimitation of a study is a systematic bias that the researcher intentionally introduces into the study design or instrument (Price & Murnan, 2013). There were two study delimitations. The first delimitation was that the influence of cognitive abilities on mathematics performance in this study relate only to the learning outcomes tested in the content domain in the 2017 May/June CSEC mathematics examination. The second delimitation was that given that the sample was stratified by high-scoring students and low-scoring students, demographic delimitations may include, age, sex, school type (private, public), school composition (single sex, co-educational), and school location (rural, urban).

Limitations

A limitation of a study design or instrument is a systematic bias that the researcher could not or did not control (Price & Murnan, 2013). The following are limitations to the study design:

- The research design is a nonexperimental, cross-sectional design and I used archival data. This design makes it difficult to make causal inferences (Bono & McNamara, 2011; Levin, 2006).
- Threats to the internal validity of the study included the reliable measure of student mathematics and cognitive abilities.

- The use of archival data eliminates the opportunity to influence how the data were captured and organized for analysis. Using this design, I observed the phenomena as it occurs naturally (see Radhakrishnan, 2013).
- Some questions on the test included more than one content domain and the scores for those content domains could not be disaggregated; as a result, I omitted these questions from the analysis.
- The number of marks allocated by cognitive domains and the content domains may not be sufficient to make a reasonable conclusion about students' abilities.
- The total number of marks assigned to the cognitive domains was not consistent across the content domains.
- The study included data from the 2017 May/June CSEC examination only. Given that the data were based on students' mathematics scores in one particular year, it is possible that a study conducted using students' mathematics scores from another year may yield different results.

Significance of the Study

Globalization and the emergent knowledge-based economy, propelled by advances in information and communication technology, have precipitated changes in the type of competencies required to function effectively in a dynamic society (Brochu, Deussing, Houme, & Chuy, 2013). Mathematics constitutes one such key competency and is considered a civic right that should be a goal for all students (Karakolidis et al., 2016; Moses & Cobb, 2001; Schoenfeld, 2002). Mathematics education comprises two dimensions: content domains, or tasks, and cognitive domain, or skills required to solve the tasks (Männamaa et al., 2012). The importance of content domains and cognitive domains has been recognized by international researchers who conducted research on both the cognitive domains and content domains in the TIMSS and the Programme for International Student Assessment (PISA; George & Robitzsch, 2018; Zhang et al., 2017). However, a search of the literature reveals a lack of research on cognitive abilities and mathematics achievement at the secondary level in the Caribbean. I sought to fill this gap and contribute to the field of research on mathematics cognitive domains and content domains by adding a Caribbean secondary perspective, targeting Vincentian students.

The minister of education in Saint Vincent and the Grenadines noted that the government understands the importance of education to the socio-economic development of Saint Vincent and the Grenadines and, hence, continues to invest in the sector (Prince, 2018). Investment in education today cost substantially more than it costs a century ago (Burnette, 2019; Psacharopoulos & Patrinos, 2018). With such sizable investments, the returns on investment and effectiveness of the education system are of great interest to governments (Cunningham, Cunningham, Halim, & Yount, 2019). The outcomes of this study may contribute to an in-depth understanding of how the performance of high-scoring Vincentian students and low-scoring Vincentian students in CSEC mathematics cognitive domains differ across the content domains. This knowledge is intended to assist educators, administrators, and teachers in Saint Vincent and the Grenadines to better target instructional areas for remediation that will support the learning and achievement of mathematics among Vincentian students. The making of policy-relevant decisions in the education sector relies heavily on the collection and use of education statistics

(Caceres, de la Peña, Di Prisco, Pineda, & Solotar, 2014). Hence, the outcome could also influence policy decisions regarding curricular changes to more directly target students' learning needs, and professional development of teachers better prepare them to deliver the curriculum in a more meaningful way, that will result in enhanced student learning. More mathematically competent persons will contribute to a more advanced society as these persons will be able to access higher paying career options and will enjoy better quality lives and contribute positively the development of the country.

Summary

A high level of mathematics proficiency is required for success at the individual and societal levels (Lipnevich et al., 2016). However, Vincentian students continue to perform poorly in the regional CSEC mathematics examination. The poor performance is evident in the low scores achieved in the cognitive domain of reasoning and across all content domains (areas). This phenomenon of poor mathematics performance is a problem for educators and policy makers in Saint Vincent and the Grenadines. A search of the literature revealed an absence of research on the influence of cognitive abilities, as defined by Bloom taxonomy (Bloom et al., 1956), on mathematics achievement at the secondary level in the Caribbean. In this research, I sought to fill that gap by providing insights into the influence of cognitive abilities on the achievement in mathematics on select content domains. I also sought to determine the extent to which the CSEC mathematics scores of high-scoring Vincentian students versus low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning differ across the content areas of algebra, geometry, measurement, statistics, and RGF. The study was quantitative in nature. I used a cross-sectional design and archival data comprising students' scores in the 2017 May/June CSEC mathematics examination. Data analysis included descriptive statistics, correlation, two-way MANOVA, and follow-up two-way ANOVA. The outcomes of the study may help to enact education reform in Saint Vincent and the Grenadines, with a view to improving students' mathematics achievement. In Chapter 2, I outline the theoretical framework for the study and discuss its application to the present study. I also synthesize the literature relating to cognitive abilities and mathematics achievement.

Chapter 2: Literature Review

Introduction

Mathematics achievement is a major component of overall academic achievement (Vista, 2016) and fundamental to the advancement of economic development, particularly in developing countries (Bosman & Schulze, 2018). Hence, educators and policy makers in Saint Vincent and the Grenadines are deeply concerned about the poor performance of Vincentian students in the CSEC mathematics examination. The poor performance is evident across all content domains and particularly on questions that require the application of higher-order thinking skills. The purpose of this study was to determine the extent to which the CSEC mathematics scores of high-scoring Vincentian students versus low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning differ across the content domains of algebra, geometry, measurement, statistics, and RFG. The review of current literature presented in this chapter includes research relevant to this study, particularly research on the influence of cognitive factors on mathematics achievement.

A systematic search of the literature is critical in unearthing studies relevant to the construct to be investigated. Consequently, this chapter begins with an outline of the literature review strategy employed, followed by a discussion of the theoretical framework underpinning the study. Following the discussion on the theoretical framework is a discussion on large-scale assessment and an outline of the development of regional mathematics assessment, CXC. I then review the five mathematics content domains investigated in the study. The review includes an explanation of the importance of each domain to overall mathematics achievement. The remainder of the chapter includes a review and synthesis of the literature relating to the role of cognitive abilities in mathematics achievement. The chapter concludes with a summary of the main findings of previous research that impact the present study, and I provide a context for the research question.

Literature Search Strategy

The literature review for this study includes a synthesis and analysis of research studies related to cognitive factors and mathematics achievement. I identified the research studies from relevant peer-reviewed articles, books, and websites relating to cognitive abilities and mathematics achievement. The databases that I searched for relevant literature relating to the study included Education Research Complete, Education Source, SAGE Journal, ERIC, Google Scholar, and ProQuest Central. I accessed the databases through the Walden University library. Keyword searches that yielded useful results included *cognitive domains, cognitive abilities, mathematics achievement, mathematics performance, mathematics content domains, mathematics strands, algebra, geometry, measurement, statistics, statistics in mathematics, statistics and mathematics curriculum, graphs, large scale assessment,* and *international large-scale assessment.* The initial searches spanned 5 years, 2015 to 2019, but due to the lack of research on this topic, I extended the search to include research studies for the last 10 years. It was also necessary to include some seminal literature, particularly regarding the theoretical framework and the mathematics content domains.

Theoretical Foundation

The theoretical base for this study was Bloom's taxonomy of educational objectives (Bloom et al., 1956). Bloom's taxonomy of educational objective is a pedagogical tool designed to guide educators in developing meaningful assessment of learning outcomes (Ramirez, 2017). The taxonomy has filled a void by providing a basis by which educators can systematically evaluate students' learning (Bertucio, 2017). Bloom's taxonomy is a taxonomy of general competence and educational objectives that provides an organized system of classifying assessment methods. The taxonomy represents a cumulative hierarchical organization of six global educational objectives (Ursani et al., 2014). The objectives, defined in behavioral terms, are knowledge, comprehension, application, analysis, synthesis, and evaluation. The hierarchy represents mental processes from simple to complex, concrete to abstract, and mastery of each simple category, which is a prerequisite to mastering the next complex category (Anderson et al., 2001; Lipscomb, 1985).

Summary of Bloom's Taxonomy

Bloom's taxonomy (Bloom et al., 1956) comprises six levels of cognitive skills, hierarchically arranged from lower-order thinking skills requiring minimal cognitive processing to higher-order thinking skills requiring deeper learning and a greater degree of cognitive processing (Adams, 2015). Figure 1 shows the hierarchical arrangement of Bloom's taxonomy, including the type of verbs used to assess each level.

Figure 1. The original Bloom's taxonomy (Bloom et al., 1956).

Knowledge. Knowledge is the foundational cognitive skill. *Knowledge* refers to "the retention of specific, discrete pieces of information including facts and definitions or methodology, such as the sequence of events in a step-by-step process" (Adams, 2015, p. 1). The knowledge objectives address predominantly the psychological process of remembering (Ramirez, 2017). Students at the knowledge level merely recall and recognize information without demonstrating an understanding of the material (Granello, 2001).

Comprehension. *Comprehension* is defined by the ability to grasp the meaning of materials. Students demonstrate comprehension by interpreting or translating material from one form to another. They display a basic understanding of the material and can summarize the main points of an article and manipulate, represent, and paraphrase information in their own words, as well as classify items into groups and compare and contrast entities (Adams, 2015; Granello, 2001).

Application. *Application* is defined as "the ability to use learned material in new and concrete situations and includes applying rules, methods, concepts, principles, and theories" (Granello, 2000, p. 4). Students at the application level can select main ideas, apply concepts and principles to new situations, apply theories to practical situations, and solve problems (Granello, 2001).

Analysis. *Analysis* refers to "the ability to break down material into its component parts, and may include the identification of the parts, analysis of the relationship between the parts, and recognition of the organizational principles involved" (Granello, 2001, pp. 4-5). Students at the analysis level can recognize unstated assumptions and logical fallacies in reasoning, distinguish between facts and inferences, and evaluate the relevancy of data (Granello, 2000).

Synthesis. *Synthesis* refers to "the ability to put parts together to form a new whole. The student originates, integrates, and combines ideas into a product, plan, or proposal that is new to him or her" (Granello, 2001, p. 297). Objectives at this level focus on creative behaviors and emphasize the formulation of new patterns or structure. Students at the synthesis level can integrate ideas from different areas into a plan to solve a problem, formulate a new schema for classifying objectives or ideas, and posing a plan for an experiment (Granello, 2000).

Evaluation. *Evaluation* refers to "the ability to judge the value of materials for a given purpose. The judgements are based on defined criteria that are either developed by the student or given to the student by an outside source" (Granello, 2001, p. 297). Evaluation is the highest level in the cognitive hierarchy— it subsumes elements of the other categories and includes conscious value judgement based on clearly defined criteria. Some of the value judgements include judging the logical consistency of written material, judging whether conclusions are adequately supported by data and applying internal and external criteria to judge one's own performance (Granello, 2000).

Bloom's taxonomy has provided a framework for systematically assessing student behavior as a result of their participating in an educational experience. The taxonomy was intended to form a universal language among teachers and assist them in creating testing materials that more accurately assess their curriculum aim (Bertucio, 2017; Hadzhikoleva, Hadzhikolev, & Kasakliev, 2019). Bloom's taxonomy is a taxonomy of general competence that serves to guide educational objectives and has been used to improve pedagogy and assessment methods in many disciplines (Ursani et al., 2014). Bloom (1956) presented his taxonomy of educational objectives in what was arguably one of the most influential education monographs of the past half century (Cullinane & Liston, 2016). It is also used as a model for identifying the cognitive processes examiners use to solve test items (Bloom et al., 1956). The taxonomy provides a useful guide to help instructors structure and sequence learning outcomes to reflect progressively difficult learning processes by providing scaffolding to help learners progress from lower levels of learning, such as knowledge and comprehension, to more cognitively demanding levels such as synthesis (Ramirez, 2017). Bloom's taxonomy does not prescribe moving from one level of objective to the next in a fixed, rigid manner; however, the progression along the continuum facilitates a logical and sequential organization of the learning process that aids mastery of the material. The taxonomy provides direction and clarity in designing the teaching process and helps instructors to be aware of the levels of difficulty of the various pedagogical activities (Ramirez, 2017). The simplicity of the taxonomy allows for clear distinction of higher-order and lower-order assessment tasks (Cullinane & Liston, 2016). Incorporating Bloom taxonomy-based objectives has been found to improve the attainment of learning outcomes (Almerico & Baker, 2004).

Bloom's taxonomy of educational objectives (Bloom et al., 1956) has provided a foundation for the understanding of learning outcomes and a platform for the development of other taxonomies, including the revised Bloom's taxonomy (RBT) by Anderson et al. (2001); Marzano's new taxonomy (MNT) by Marzano and Kendall (2006); and mathematical wellbeing (MWB) by Clarkson, Bishop, and Seahs (as cited in Irvine, 2017). Despite the development of more recent taxonomies, the original Bloom's taxonomy has been used extensively by educators to identify and delineate tasks involving higher-order and lower-order thinking skills (Cullinane & Liston, 2016; Irvine, 2017). Bloom's taxonomy has also been found to be exceptionally helpful in providing clarity in designing the teaching process by structuring and sequencing educational objectives in needs assessment, lesson planning, and assessment (Ramirez, 2017).

Bloom's taxonomy of educational objectives (Bloom et al., 1956) has influenced assessment at all levels, from classroom tests to international large-scale assessments, including TIMSS (Abu Tayeh, Mohammad, & Mohammad, 2018; George & Robitzsch, 2018; Mullis & Martin, 2019). These test designs begin with a table of specifications which is usually a 2-way matrix that specifies the content domains and cognitive abilities to be tested. The table of specifications provides a guideline for obtaining a representative sample of test items (Gierl, 1997). TIMSS's mathematics assessment is modelled from

Bloom's taxonomy of educational objectives and uses three levels of cognitive domains, namely, knowing, applying, and reasoning, in its fourth- and eighth-grade mathematics assessment. The cognitive domains are used to develop a content-by-process matrix to create mathematics assessment in the content domains of number, algebra, geometry, and data and probability for eighth-grade students internationally. The CXC has also adopted Bloom's taxonomy in the creation of its regional Grade 11 CSEC mathematics examination. The CSEC mathematics assessment follows a similar content-by-process matrix in TIMSS's Grade 8 mathematics assessment (Martin & Mullis, 2019). The CSEC mathematics assessment includes content domains of computation, number theory, consumer arithmetic, sets, measurement, statistics, algebra, relations, functions and graphs, and geometry and trigonometry at the cognitive domains of knowledge, comprehension, and reasoning. Table 3 shows the comparison of Bloom's taxonomy with TIMSS's and CXC's cognitive domains.

Table 3

Cognitive Domains: Bloom's Taxonomy, TIMSS, and CXC

Bloom's taxonomy	TIMSS cognitive domain	CXC cognitive domain
Knowledge - Recall of information, methods, procedures, pattern, structure, and settings	Knowing - Covers facts concepts, and procedures	Knowledge – Recall of rules, procedures, definition and facts
Comprehension - Understand assessment material, and translate it into own words	Applying - Focusses on the ability to apply knowledge and conceptual understanding to solve problems or answer questions	Comprehension - The use of algorithms and the application of these algorithms to familiar problem situations
Application - Apply knowledge to new situations	Reasoning - Encompasses unfamiliar situations, complex contexts and multistep problems	Reasoning - Solving nonroutine problems, making inferences and generalizations, analyzing and synthesizing

The original Bloom's taxonomy (Bloom et al., 1956) has provided educators and instructional psychologists with a framework for designing instructions to capitalize on the way students learn. However, after almost 5 decades of using Bloom's taxonomy, many educators questioned the validity of the taxonomy for meeting the needs of students and educators (Darwazeh, 2017). This concern led Anderson et al. (2001) to revise and update Bloom's taxonomy to make it more relevant to the needs of the 21st-century students and teachers (Anderson et al., 2001). The changes to the taxonomy were in relation to the terminology, structure, and emphasis (Forehand, 2010; Krathwohl, 2002). Although the original Bloom's taxonomy used nouns to describe the levels, the revised taxonomy used verbs. Three categories were renamed from knowledge to remember, from comprehension to understanding, and from synthesis to create. The evaluation objective was changed to evaluate and placed as the penultimate category, whereas synthesis was renamed create and place as the highest level (Darwazeh, 2017).
Figure 2 shows the comparison of the original Bloom's taxonomy (Bloom et al., 1956) and the revised Bloom's taxonomy (Anderson et al., 2001).

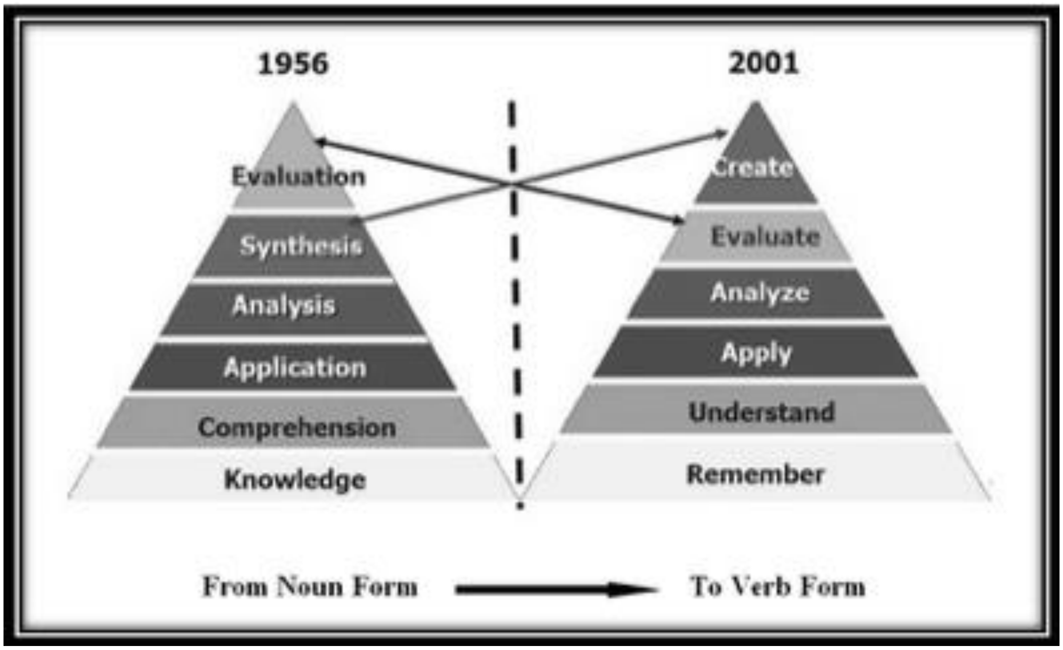


Figure 2. The revised Bloom's taxonomy by Anderson et al. (2001).

Saint Vincent and the Grenadines' Demographic Information

Country Profile

Saint Vincent and the Grenadines is a chain of 32 islands and cays (nine of which are inhabited) in the Caribbean, with a total area of 389 square kilometers or 150 square miles. Saint Vincent and the Grenadines is a volcanic, mountainous island. The climate is tropical. Its capital is Kingstown and the other towns are Calliaqua, Chateaubelair, Georgetown, Layou, and Barrouallie. The main income earner is agriculture. Saint Vincent and the Grenadines gained independence from Britain in 1979 (Fraiser, 2019). Currently, there are two main political parties. The population count as of January 01, 2019, was 109,545, and is predominantly African Blacks which accounts for (66%), Mixed (19%), East Indian (6%), European (4%), Carib Amerindian (2%) and other (3%) (United Nations World Population Prospects, 2019). The main language is English, however, there is also creole English and French patois. Saint Vincent and the Grenadines is a Christian country, the main religions include Anglican, Catholic, Methodist, Pentecostal, and Spiritual Baptist.

The Education System

The education system in Saint Vincent and the Grenadines comprises three levels–primary (7 years), two-phased secondary education (the first phase is 5 years, and second phase is 2 years), and tertiary (Education Act, 2006). There are 43 primary school, 26 secondary schools, and one tertiary institution comprising various divisions. Secondary education usually spans Ages 11 to 18. Of the 26 secondary schools, 20 are government owned, and six are government assisted. There are two single sex male secondary school, and two single sex female secondary schools. Seven of the 26 secondary schools are located in the urban area, 17 in the rural area, and two in the Grenadines. Education at the lower secondary level is modelled on a national curriculum, whereas education at the upper secondary school is dictated by the syllabi developed and examined at the CSEC level by the CXC. Students who are not in a formal school setting may also write the CSEC examinations as private candidates. In 2017, 1,713 students from 26 Secondary schools and 10 private institutions wrote the CSEC mathematics examination.

Literature Review Related to Key Concepts

Large Scale Assessment

Large-scale assessment (LSA) is a summative assessment, or an ‘assessment of learning’ (Klieger, 2016). It is a tool used for educational accountability (Copp, 2017;

Klinger, DeLucas, & Miller, 2008). LSA of student achievement reveals how students perform on literacies, and the types and levels of achievement in relation to correlates of learning, including student background, attitudes, and perceptions, as well as home and school characteristics (Anderson, Lin, Treagust, Ross, & Yore, 2007). The intent of large-scale assessment is to measure learning outcomes for accountability purposes (Cox & Meckes, 2016; Decker & Bolt, 2008; Klieger, 2016; & Looney, 2011). Large-scale assessment aims to promote student achievement by holding educators accountable (Decker & Bolt, 2008; Klinger et al., 2008; Miller, 2013).

International large-scale assessments (ILSAs) began in 1958 by the UNESCO institute for education (Husén, 1979). The impetus for ILSAs is to study the educational achievement and its determinants in different countries by collecting reliable, valid, and comparable information about student abilities and analyzing this information to better understand the relationship among student abilities and educational, social, and economic phenomena (Yamamoto & Lennon, 2017). Countries can use the results from ILSA to learn from each other, and avoid pitfalls (Johansson, 2016). ILSA allows for comparative evaluation of the education system of countries, thereby revealing gaps between first world nations and high-income countries (Cox & Meckes, 2016).

ILSA include the First International Mathematics Study (FIMS), Trends in, Mathematics and Science Study (TIMSS) and the Progress in International Reading Literacy Study (PIRLS), conducted by the International Association for the Evaluation of Educational Achievement (IEA). There is also, the Program for International Student

Assessment (PISA), conducted by the Organization for Economic Cooperation and Development [OECD] (Sui Chu Ho, 2016).

The two most popular international large-scale mathematics assessment are TIMSS and PISA. TIMSS is an integrated assessment of mathematics and science conducted at fourth and eighth grade levels. The assessment is designed to measure trends in student performance. The assessment was first conducted in 1995, and subsequently 4 years thereafter. Over 55 countries participate in TIMSS, representing a wide range of geographical and economic diversity (Mullis, Martin, Foy, & Hooper, 2015). TIMSS collects information about the students, teachers and classroom characteristics, and includes teacher and student background and experiences. This information provides a context in which the results are reported (Martin, Mullis, & Hooper, 2015). Comparison of students' performance is benchmarked internationally (Balázsi & Szepesi, 2018). TIMSS mathematics assessments at both the fourth-grade and eighth-grade levels are organized around two dimensions:

- the content dimension, which specifies the subject matter to be assessed, and
- the cognitive dimension, which specifies the thinking processes to be assessed (Lindquist, Philpot, Mullis, & Cotter, 2019).

The content domains assessed at each grade level differ, reflecting the mathematics that is taught at the respective levels. However, the same cognitive domains are assessed at fourth-grade and eighth-grade, but with a shift in emphasis. The assessment includes a range of problem-solving situations within mathematics, with emphasis on the higher-order thinking skills, such as applying and reasoning (Lindquist et al., 2019).

PISA is an OECD sponsored program which aims to evaluate the education system of countries by testing the competencies and skills of 15-year old students at the end of compulsory schooling (Ninomiya, 2019; OECD, 2014; Rautalin & Alasuutari, 2009; Yalçin & Tavşancil, 2014). The program was initiated in 2000 and currently includes 90 countries. Students are assessed triennially in three main literacies: reading, mathematics and science (Lewis, 2017; OECD, 2014). In its triennial assessment, PISA focusses on 1 literacy proficiency from among the three domains. In addition to the three subject domains, PISA includes innovative domains such as collaborative problem solving, global competitiveness, and financial literacy (OECD, 2016). The focus of PISA's assessments is on the application of knowledge learned in school to real-life situations. PISA's mathematics literacy is defined as

> an individual's capacity to identify and understand the role that mathematics plays in the world, to make well-founded judgements, and to use and engage with mathematics in ways that meet the needs of that individual's life as a constructive, concerned, and reflective citizen (Anderson et al., 2007, p. 593)

Like TIMSS, in addition to assessments in subject domains, PISA also evaluates the socio-economic indicators of students and their parents as well as the school environment–how the school is managed. Socio-economic indicators include student background information, such as home resources and parents' occupation and level of education (Anderson et al., 2007; İnce & Gözütok, 2018; OECD, 2016).

International large-scale assessments such as TIMSS and PISA play a significant role in contemporary educational landscape by influencing education reform policies at the global, regional and international levels with a view to improving educational practices and performance (Elliott, Stankov, Lee, & Beckmann, 2019; Grek, 2013; Ninomiya, 2019; Ozga, 2012). Many countries overhaul their education system following the release of the results of TIMSS and PISA assessments. This was the case of Japan after the publication of the PISA 2003 results (OECD, 2004). The Japanese government refocused school teaching and curriculum from an emphasis on 'solid academic ability' to 'PISA-style Literacy' (Ninomiya, 2019). According to Matsushita, and Oohashi (as cited in Ninomiya, 2019), ILSAs have also influenced the creation of an 'evidence-based improvement cycle' and a corresponding 'target management system' as part of Japan's education reform. ILSAs have also influenced education reform in Spain where Spanish students reportedly performed poorly on TIMSS, PIRLS, and PISA. The poor performance of students on the 2012 PISA assessment led the government of Spain to undertake education reform that specifically targeted the secondary level (Choi & Jerrim, 2016). In Italy, PISA's results have influenced the implementation of initiatives at the local and national levels, including teacher training and retooling, and delivery, and support to schools aimed at reducing education disparities between the rich and poor. The initiatives include the teaching of the 3 subject domains assessed by PISA, targeting areas of deficiencies in student ability (Damiani, 2016).

The assessments administered by PISA and TIMSS serve as ranking for comparison and benchmarking tools for student achievement at the end of compulsory education (Güvendir, 2017; Sahlberg, 2011; Želvys, 2017). For instance, PISA's benchmark for 2022 requires students to achieve at least the 3rd-level in reading with no less than 49%. Benchmarks for mathematics and science are set at 51% and 56% respectively. ILSAs respond to the global education reform movement by providing opportunities to compare the achievement of 15-year old students in various countries within a common education space (Želvys, 2017). ILSAs provide objective and global evidence of the comparative effectiveness of the education systems in participating countries (Adamson, Forestier, Morris, & Han, 2017). The ranking data provided by both TIMSS and PISA have inspired lower performing countries to seek educational best practices from the better performing countries. Policymakers from England have led fact finding missions to Hong Kong, Shanghai and Singapore in search of a formula for success and a system on which to model their pedagogical practices (Adamson et al., 2017; You, 2018).

ILSAs have created a sagacity of global educational accountability, thereby influencing the emphasis on national accountability mechanisms (Breakspear, 2012; Sellar & Lingard, 2014). ILSAs encourage accountability and include systems accountability at all levels, including students, teachers, schools and districts levels. System accountability can carry high stakes for schools when assessment results are used to streamline or reconstitute underperforming schools or districts (Goertz & Duffy, 2003). LSA, at any level, is accompanied by many consequences including high stakes for schools or districts, teachers and students (Decker & Bolt, 2008). High-stakes consequences of LSAs for schools or districts is evident when the assessment outcomes are used to restructure or reconstitute under-performing schools or districts (Goertz & Duffy, 2003). High-stakes consequences of LSAs for teachers is apparent when the assessment results are used to influence decisions regarding teachers' evaluation, performance pay and continued employment (Braden, 2002). High-takes consequences of LSAs for students is ostensive when assessment results are used to determine a student's fate, such as whether a student is promoted or retained at a grade level, whether a student will graduate or be assigned to a particular class, program or school (Goertz & Duffy, 2003).

There are both beneficial and detrimental effects associated with LSAs. LSAs provide a common assessment or 'yardstick' by which all students taking the assessment are measured, thereby ensuring that all students are treated fairly and equitably (O'Connor, 2017; Phelps, 2012). LSAs have also been found to have a positive effect on student achievement and to promote curricular alignment among state standards, classroom assessment and tests (Decker & Bolt, 2008; Phelps, 2012). LSAs also promote equity among traditionally at-risk groups of students, or students with special needs (Decker & Bolt, 2008; Roderick & Engel, 2001). LSAs follow rigorous test administration and result in sizeable data sets which provide opportunities for investigating pedagogy and classroom practice (Heyneman & Lee, 2013; Howie & Plomp, 2006; Wagemaker, 2013). The data sets from LSAs allow for secondary analysis by scholars. Researchers can measure achievement trends within countries, as well as engage in evidence-based enquiry (Gustafsson, 2008; Johansson, 2016).

Notwithstanding the benefits, LSAs have been criticized for causing a narrowing of the curriculum, encouraging an emphasis on lower-order thinking skills at the expense of higher-order thinking skills, and reducing instructional time at the expense of test preparation activities. It is also believed that LSAs encourage ‘teaching to the test’, and neglects content not covered in the assessment, rather than fostering the acquisition of general knowledge and skills (Decker & Bolt, 2008; Johansson, 2016; Rogers, 2014). LSAs provide rich data on student achievement, however, Rutkowski and Delandshere (2016) caution against making causal claims regarding student achievement, particularly where the mechanisms or causal explanations for a phenomenon may not be comparable across contexts such as groups and counties. The high-stakes accountability associated with LSAs have been thought to encourage cheating and reduce professionalism among teachers (Chester, 2005a; Chester, 2005b; Cizek, van der Linden, & Cook, 2012; Shephard, 2010). Other negative consequences associated with accountability in LSAs include questionable evaluation of teachers, resulting in increased teacher stress, and unwarranted reduction in teacher salaries and school sanctions (Rogers, 2014).

Mathematics Content Domain

Student success in mathematics requires mastery of key foundation mathematics concepts. Mathematics content is a body of knowledge organized by domains which provides a structured approach to the learning of mathematics. Some mathematics domains represent core content and concepts which are the foundation of mathematics and which act as the gateway for learning higher mathematics. Mastery of basic mathematics content provides the impetus for learning more advanced mathematics content. The 5 mathematics content domains addressed in the study, are algebra, geometry, measurement, statistics, and RFG.

Algebra. *Algebra* has been described as the foundation of mathematics (Ferrini-Mundy, 2000; Greeno & Collins, 2008; Lee, Ng, & Bull, 2018; Litke, 2020a; Litke, 2020b; MacGregor, 2004). Algebra is used in many phases of life, including solving everyday problems. It provides the tools to represent and analyze quantitative relationships (Knuth, Stephens, Blanton, & Gardiner, 2016). Algebra is therefore considered a gateway to future educational and occupational opportunities (Pedersen, 2015). Schoenfeld (1995) likened the importance of algebra in the 20th century to reading and writing in the industrial age. Algebra is essential for understanding science, statistics, and business as well as functioning in a technological environment (Schoenfeld, 1995). Schoenfeld further posited:

> Algebra has become an academic passport for passage into virtually every avenue of the job market and every street of schooling. With a few exceptions, students who do not study algebra therefore are relegated to menial jobs and are unable often even to undertake training programs for jobs in which they might be interested. They are sorted out of the opportunities to become productive citizens in society. (pp. 11–12)

In contemporary society, algebra is considered a pivotal concierge to higher-level mathematics and a predictor of future academic success (Prendergast & Treacy, 2018). Given such prominence, algebra is assessed in many large-scale assessments including TIMSS, PISA, and National Assessment of Education Progress at the international level and CSEC at the regional level. Algebra accounts for 30% of the content domain in TIMSS's 8th-grade mathematics assessment (Mullis & Martin, 2019) and 26% in CSEC mathematics (CSEC, 2008).

Recent researchers focused on various aspects of algebra including the teaching and learning of algebra at different levels of education. Some researchers focused on intervention strategies used by teachers to improve the learning of algebra (Cohen, 2018; Litke, 2020a; Litke, 2020b; Prendergast & Treacy, 2018; Rau & Mathews, 2017; Stylianou et al., 2019). Other Researcher investigated the optimal time for introducing students to algebra (Lee et al., 2018). Other researchers, including Barbieri, Miller-Cotto, and Booth (2019) examined the type of errors students make when solving algebraic problems. The researchers found that students' general misconceptions about mathematics affected their algebraic problem-solving abilities as evident in the high number of conceptual errors demonstrated. The role of cognitive abilities in the learning of algebra was explored by Roegner (2013) who found that university students who rely on lower-order thinking processes, such as procedural approaches were least successful in solving algebraic problems than their counterparts who adopted a conceptual approach.

Geometry. *Geometry* is considered one of the most important branches of mathematics (Ünlü & Ertekin, 2017). Crompton et al. (2018) defined geometry as "the study of properties, relationships, and transformations of spatial objects, within an interconnected network of concepts and representational systems" (p. 59). Geometry is fundamental to many aspects of everyday life (Cass, Cates, Smith, & Jackson, 2003). Geometry is important for understanding space. It provides students with a foundation for understanding other areas in mathematics (Galitskaya & Drigas, 2020). Geometry is used to explore the characteristics and relationships of angles, lines, and shapes (Üstün & Ubuz, 2004). Success in mathematics is dependent on a sound understanding of geometry concepts (Education Review Office, 2018). Knowledge of geometry helps to develop students' decision-making and judgement skills, and provides them with a foundation for advanced mathematical subjects, particularly in the area of STEM (Zhang, Ding, Stegall, & Mo, 2012).

Geometry is a critical component of mathematics (Jiang, Li, Xu, & Chen, 2019). The geometry assessment strand is an essential mathematics strand is essential in other facets of mathematics (Ferrini-Mundy, 2000). Geometry was developed from the practical needs of daily life and influences a number of other disciplines, including natural sciences, and social studies (Atasoy, 2019; Ferrini-Mundy, 2000; Kilicoglu, 2020) and careers, such as art, architecture, and engineering (Ferrini-Mundy, 2000). Geometry is used to solve problems in other areas in mathematics, including measurement (Kilicoglu, 2020; Syarifudin, Purwanto, Irawan, Sulandra, & Fikriyah, 2019), as well as promoting understanding in other areas of mathematics, including number and operations, measurement, algebra, data analysis, and probability (Ferrini-Mundy, 2000). As an everyday language, geometry helps in describing places such as 'parallel to' and 'adjacent to'. Geometry is used to create an appreciation for the beauty of nature, by providing a way to interpret the physical environment, hence enhancing students' reasoning and justification skills (Ferrini-Mundy, 2000). Student geometric knowledge has been found to be related to mathematics achievement, as well as overall academic achievement (Giofrè, Mammarella, & Cornoldi, 2014; Giofrè, Mammarella, Ronconi, & Cornoldi, 2013).

Measurement. *Measurement* is one of the foundation concepts in mathematics. Knowledge of measurement is required for day-to-day functioning in the world (Hurrell, 2015). A comprehensive understanding of measurement is critical in the STEM field (Doabler et al., 2019). In the field of engineering, measurement is used to obtain precise estimates of mass and strain. Epistemologists also use measurement to determine trends in health-related events (Paules, Marston, & Fauci, 2019). Measurement is a very important education objective from kindergarten through the elementary years (Castle & Needham, 2007). Given the ubiquitous nature of measurement, Serow, Callingham, and Muir (2014) postulate that persons who lack knowledge of measurement, lack the capacity to effectively and efficiently operate in society, both personally and professionally, could not be considered numerate. Students with a lack of understanding of measurement are not likely to achieve overall mathematics proficiency (Doabler et al., 2019). Measurement provides a context for learning other mathematics concepts, including place value, number, geometry, and probability (Van de Walle, Karp, & Bay-Williams, 2013). Students are able to relate to measurement as they can often see its usefulness and can relate many of the tasks in measurement to their daily lives (Reys et al., 2012).

Measurement is a powerful strand of mathematics. It has rich pedagogical possibilities and can create opportunities for rigorous and meaningful learning of mathematics. The teaching of measurement integrates well with other curricular subjects, such as science, geography, music and history (Hurrell, 2015). In underscoring the importance of measurement in the curriculum, Reys et al. (2012) posit that the importance of measurement does not relate so much to the mathematics, but more so the pedagogical benefit. They summarize the importance of measurement as an effective way to engage and motivate students who would not normally be motivated to learn other topics.

Statistics. *Statistics*, as a branch of mathematics, is essential for functioning in a society that is becoming more data-driven and digital. Statistics fosters critical thinking skills, and as such, the general population should be exposed to a basic understanding of statistics (Capaldi, 2019). A thorough understanding of statistics is required in the STEM field (Paules et al., 2019), as well as in achieving mathematics proficiency (Doabler et al., 2019). We use statistics in every facet of our lives, often without being conscious (Spiegelhalter, 2020). Given its importance, statistics has gained prominence at all levels of education. According to Goldstein (2007), statistics should form a central feature in the mathematics curriculum. Statistics range from the simple throwing of a dice to statistical investigations including data collection, data representation (which involves the production of graphs and tables), and data reduction (finding means and ranges, and drawing inferences). The study of statistics, or chance and data, is as important as the study of algebra, and is essential in the training of students as future citizens (Callingham & Watson, 2017; Watson, 2001). Statistics is an area of applied mathematics which can be considered an appropriate vehicle for motivating students to learn mathematics. It incorporates a wide variety of interpretative and manipulative skills and provides opportunities for students to apply these skills to other areas, including arithmetic and graphs. Many students can relate to statistics as it presents real-life scenarios and familiar problems relating to other areas of study (Goldstein, 2007).

In accentuating the importance of statistics as a strand of mathematics, Goldstein (2007) wrote:

> A basic understanding of statistical ideas, and especially the idea of statistical modelling involving exposure to statistical data analysis, is as fundamental to an understanding of modern society and its artifacts as is language literacy. From this, it follows that statistical knowledge and practice should suffuse the school curriculum. (p. 8)

Goldstein argued for the retention of statistics as an integral part of the schools' curriculum for the foreseeable future. Goldstein's argument is consistent with the view of Mills (2004), that developing statistical thinking and reasoning skills are important objectives in society. Moore (2007) advocated establishing synergy between content, pedagogy, and technology in the teaching of statistics. According to Moore (2007), the nature of statistic lends itself to the active participation of students in the learning process and should be extended to include non-mathematical statistical concepts and ideas.

Graphs. Graphing is an important curriculum area in preparing students for 21st-century careers in STEM (Larson & Whitin, 2010; NCTM, 2006; STEM Education Coalition, 2009). Graphing is usually associated with collecting and interpreting numerical information and is deemed a vital skill in a world that is inundated by data. Graphing also provides significant opportunities for students to represent and communicate important mathematical relationships (Larson & Whitin, 2010). The construction and interpretation of graphs are essential mathematics activities (Ellis, Tasova, & Singleton, 2018). The study of graphs affords students opportunities to integrate mathematics with other areas of learning. An understanding of graphs is critical to understanding chance and data, as well as working mathematically and scientifically, including investigating and communicating, and in general, participating effectively in society and the environment (Lake & Kemp, 2001). Knowledge of graphs is important to interpreting scientific factors, analyzing data, and analyzing patterns (Berg & Boote, 2017). Understanding graphs is considered a higher order thinking skill as it includes reading, interpreting, and synthesizing information represented in various pictorial forms (Patahuddin & Lowrie, 2019).

CSEC Mathematics Cognitive Domain

The cognitive domains used in the CSEC mathematics examinations are adopted from Bloom's original taxonomy of educational objectives (Bloom et al., 1956). The first two levels, knowledge and comprehension, are used as defined by Bloom's taxonomy. The third level, application, was renamed reasoning, in the CSEC mathematics syllabus (CSEC, 2008). The CXC has defined the cognitive domains as following:

Knowledge. Items that require the recall of rules, procedures, definition, and facts, that is, items characterized by rote memory as well as simple computation, computation in measurement, construction and drawings.

Comprehension. Items that require algorithmic thinking that involve translation from one mathematical mode to another. Use of algorithms and the application of these algorithms to familiar problems situations.

Reasoning. Items that require:

(i) Translation of non-routine problem into mathematical symbols and then choosing suitable algorithms to solve the problems;

(ii) Combination of two or more algorithms to solve problems;

(iii) Use of an algorithm or part of an algorithm, in a reverse order, to solve a problem;

(iv) The making of inferences and generalizations from given data;

(v) Justification of results or statement;

(vi) Analyzing and synthesizing (CSEC, 2008).

The Importance of Mathematics

Mathematics competency is an important component of STEM and is critical to our daily lives and the success of an economy (Algarni, 2018; Hassan, Abdullah, Ismail, Suhud, & Hamzah, 2019; Panizzon et al., 2018; Primi et al., 2020; Waxman, 2020). Mathematics education helps students to develop their own knowledge and become active learners by equipping them with the resources and opportunities to explore, investigate, and make sense of real-world situations, thereby constructing a solid foundation for future success (Hassan et al., 2019). Mathematics and science literacy help to liberalize and stabilize society as well as contribute to societal development, and give citizens hope for the future (Bosman & Schulze, 2018). As an *a priori* discipline, mathematics provides science with concepts, theories and techniques for interpreting and explaining the physical world (Waxman, 2020). Mathematics plays a pivotal role in our daily lives, not only for personal success, but it is indispensable in the pursuit of careers that are important for a country's economic growth and development (Naidoo & Kapofu, 2020; Primi et al. (2010). Persons who lack mathematical competence are likely to be economically disadvantaged (Lipnevich et al., 2016). Now, more than ever, mathematics is a central part of life and is critical to making informed decisions and existing as productive citizens (Algarni, 2018). Mathematics competency is critical for success in our high-paced 21st century (Karakolidis et al., 2016). The importance of mathematics is also recognized by students who reported that though challenging, mathematics is important for future careers, especially in the field of STEM (Dobie, 2019)

Mathematics Achievement

Students' mathematics achievement is a major component of their overall academic achievement and is considered indispensable to life (Ajello, Caponera, & Palmerio, 2018; Soni & Kumari, 2017; Vista, 2016). Mathematics competence is an essential prerequisite for lifelong learning and active participation in society and culture (Ehmke, van den Ham, Sälzer, Heine, & Prenzel, 2020). Despite such high value associated with mathematics, students at all levels continue to underperform in the subject, thereby attracting attention locally, regionally, and internationally. ILSAs such as TIMSS and PISA have established benchmarks for mathematics achievement internationally. The results consistently show many countries performing below the benchmarks established for TIMSS fourth grade and eighth grade mathematics (Mullis, Martin, Foy, & Arora, 2012; Mullis et al., 2015). The 2015 TIMSS mathematics assessment included 49 countries. Tables 4 and 5 show the percentage of students achieving the various benchmarks at the fourth grade level and eighth grade level respectively.

Table 4

Percentage of Fourth-Grade Students Achieving Benchmark

Benchmark	Definition	Percentage
Intermediate	Students can apply basic mathematical knowledge in simple situations	75%
High	Students can apply knowledge and understanding to solve problems	65%
Advanced	Students can apply knowledge and understanding in a variety of relatively complex situations and explain their reasoning	6%

Note. Percentages are based on benchmark categories and do not add to 100% as they do not include scores that did not meet these benchmarks

At the fourth-grade level, only 6% of the students in the 49 countries achieved the advanced benchmark. Fourteen countries showed relative weakness in numbers, 21 in geometric shapes and measurement and 20 in data display (Mullis et al., 2015).

Table 5

Percentage of Eighth-Grade Students Achieving Benchmark

Benchmark	Definition	Percentage
Intermediate	Students can apply basic mathematical knowledge in a variety of situations	62%
High	Students can apply knowledge and understanding in a variety of relatively complex situations	26%
Advanced	Students can apply and reason in a variety of problem situations, solve linear equations and make generalizations	5%

Note. Percentages are based on benchmark categories and do not add to 100% as they do not include scores that did not meet these benchmarks.

Overall, only 5% of the students in the 49 countries achieved the advanced benchmark at the eighth grade level (Mullis et al., 2015). Twelve countries showed relative weakness in number, 14 in algebra, 19 in geometry, and 22 in data and chance (Mullis et al., 2015).

Ehmke et al. (2020) investigated the concordance between students' scores in the 2012 PISA mathematics assessment, used to define international benchmarking, and their scores in the National Education Panel Study (NEPS) in Germany. The results showed that the total sample, as well as subgroups, there were almost identical distributions of the PISA proficiency levels. The outcomes of the study provide evidence of concordant score distribution, thereby supporting the validity of the PISA benchmarks for international mathematics achievement. The results also provide insights of how national assessment could be related to international assessments (Ehmke et al., 2020).

Mathematics is a multidimensional construct that includes different cognitive skills (Gilmore et al., 2018; Männamaa et al., 2012; Ölmez, 2020). The development of mathematics skills is a complex process which requires the mastery of several subskills (Locuniak & Jordan, 2008; VanDerHeyden & Burns, 2009) and the use of various cognitive abilities (Taub, Floyd, Keith, & McGrew, 2008). A variety of nomenclature have been used by researchers to describe cognitive abilities. The Cattell-Horn-Carroll (CHC) theory of human cognitive abilities describes cognitive abilities in terms of three strata of intelligence, namely: general intelligence (*g*), broad cognitive abilities, and narrow cognitive abilities. The broad cognitive abilities include: fluid reasoning (Gf), comprehension-knowledge (Gc), short term memory (Gsm), visual processing (Gv), auditory processing (Ga), long-term retrieval (Glr), processing speed (Gs), decision/reaction time or speed (Gt), reading and writing (Grt), and quantitative knowledge (Gq). These broad cognitive abilities subsume approximately 70 narrow cognitive abilities (Floyd, Evans, & McGrew, 2003; McGrew, LaForte, & Schrank, 2014)

The CHC theory of cognitive abilities has provided a rich theoretical base for understanding human cognitive abilities and their relationships with various academic outcomes (Floyd et al., 2003). Many researchers have reported strong relationships between cognitive abilities and mathematics achievement. Achievement in mathematics is usually differentiated into two dimensions: the content of the task, which includes the topics, and the cognitive abilities needed for solving these tasks, such as knowing, computing, knowing and using algorithms, solving word problems, as well as applying these skills in novel situations (Männamaa et al., 2012). According to Gilmore et al. (2018), a thorough understanding of mathematics achievement requires an identification of important relationships between cognitive skills and specific components of mathematics.

Both cognitive and non-cognitive factors have been found to play a significant role in student mathematics achievement (Lee & Stankov, 2018; Semeraro, Giofrè, Coppola, Lucangeli, & Cassibba, (2020). However, within recent times, the role of cognitive abilities in mathematics achievement has attracted the attention of many researchers who have generally found strong associations between these cognitive abilities and mathematics achievement among students of varying ages (Areepattamannil & Caleon, 2013; Caemmerer, Maddocks, Keith, & Reynolds, 2018; Cormier, Bulut, McGrew, & Singh, 2017; Cowan, Hurry, & Midouhas, 2018; O'Connell, 2018). The influence of general intelligence and cognitive abilities on the mathematics achievement of 5-year-old to 19-year-old students was investigated using Woodcook-Johnson's (WJ) III tests of cognitive abilities (WJ COG) as the measure of student achievement (Cormier et al., 2017; Floyd et al., 2003; Giofrè, Borella, & Mammarella, 2017; Taub et al., 2008; Tolar, Fuchs, Fletcher, Fuchs, & Hamlett, 2016). Taub et al. (2008), Cormier et al. (2017), and Giofrè et al. (2017) used structural equation modelling; which includes factor analysis and multiple regression analysis to analyze the structural relationship between latent structures and measured variables while Floyd et al. (2003) used multiple regression analysis to investigate the relationship among the variables. The results of the studies were generally consistent. They all reported moderate to strong relationship between cognitive abilities and mathematics achievement. However, the strength of the relationship seemed to vary depending on the cognitive ability, the age of the student, and the type of mathematics task. For instance, both Floyd et al. (2003) and Cormier et al. (2017) found processing speed to have a moderate relationship with mathematics reasoning and a moderate to strong relationship with mathematics calculation skills during elementary years. However, Taub et al. (2008) found the relationship between processing speed and mathematics achievement to be significant. In the latter years, comprehension-knowledge (Gc) was found to be moderately related to mathematics calculation skills and moderately to strongly related to mathematics reasoning (Floyd et al., 2003). Fluid reasoning (Gf), short-term memory (Gsm), and working memory generally demonstrated moderate relations with the mathematics domains.

Fluid intelligence, also referred to as fluid reasoning (Gf), is a broad cognitive ability that has been found to play a critical role in students mathematics achievement. Primi et al. (2010) investigated the association of fluid intelligence and inter-individual differences in intra-individual growth on mathematics achievement among 13-year-old and 15- year-old students in the United States. The mathematics domains investigated were: geometry, numbers, equations, statistics, functions, and graphs. The cognitive domains were numerical reasoning, verbal reasoning, spatial reasoning, and abstract reasoning. The study found fluid intelligence to be strongly related to mathematics achievement at all ages. Students with higher fluid intelligence showed faster increases in mathematics scores than their counterparts with lower fluid intelligence. Fluid intelligence was also found to be associated with students' reasoning and problem-solving abilities. Green, Bunge, Chiongbian, Barrow, and Ferrer (2017) investigated the role of fluid reasoning in mathematics achievement among a sample of students from age 6 to 21 also in the United States. The researchers used structural equation modelling to examine the direct and indirect relations between children's previous cognitive abilities and their future mathematics achievement. Like Primi et al. (2010), Green et al. (2017) found fluid reasoning to be the only significant predictor of future mathematics for students in both primary and secondary school. Similar findings were reported by Gelbart (2007) who investigated the relationship among cognitive functioning, as defined by the CHC theory and mathematics achievement among a sample of high school students. Fluid reasoning was found to be a strong and specific predictor of mathematics reasoning. The results of these studies were further corroborated by the findings of a meta-analysis conducted by Peng, Wang, Wang, and Lin (2019), in which they sought to determine the relationship between fluid intelligence and reading and mathematics. Fluid intelligence and reading and mathematics were found to have a reciprocal relationship. However, the relationship between fluid intelligence and mathematics was stronger than that between fluid intelligence and reading, and increased with the complexity of the tasks and the age of the students (Peng et al., 2019). The findings of these studies support Cattell's conceptualization of fluid reasoning as a precursor to the development of the mathematics problem solving skills. The studies included a wide spectrum of learners, ranging from kindergarten to university, and covered a wide range of time, in some cases, over a decade apart, and ultimately the outcomes were consistent across age and time.

Working memory (wm) is another cognitive ability that is associated with mathematics achievement. Research in this area have spanned kindergarten to university. Lee and Bull (2016) investigated the relationship between working memory updating and mathematics performance from kindergarten to ninth grade and the extent to which earlier capacities in working memory updating and mathematics contributed to later development. They found that students' working memory updating capacity consistently predicted subsequent mathematics performance and that students with higher working memory or updating capacity performed better in mathematics than their counterparts with lower working memory updating capacity (Lee & Bull, 2016). The findings of Lee and Bull (2016) were supported by those of Gimbert, Camos, Gentaz, and Mazens (2019) who found working memory to be a significant predictor of mathematics achievement among 7-year-old students. Musso, Boekaerts, Segers and Cascallar (2019) analyzed the relationship between working memory capacity, executive attention, self-regulated learning, item characteristics and mathematics performance among of sample of university students (ages 18–27). The mathematics test consisted of multiple-choice items testing arithmetic, percentages, proportion, decimals, algebra and geometry. The finding of the study indicated a direct relationship between working memory capacity and mathematics performance. These findings are consistent with findings from a meta-analysis conducted by Peng, Namkung, Barnes, and Sun (2016) to determine the relationship between mathematics and working memory. They found a significant moderate relationship between mathematics and working memory. This relationship was significantly affected by the type of mathematics skills. Problem-solving tasks involving worded problems and whole number calculations showed the strongest relation with working memory, whereas geometry showed the weakest relation with working memory. The relation between working memory and algebra was moderate. Donati, Meaburn, and Dumontheil (2019) investigated the effect of working memory, inhibitory control, and processing speed on achievement in English, math, and science during adolescence. The mathematics assessment included conceptual understanding, mathematical reasoning, and problem solving. Donati et al. (2019) found working memory, reasoning, and slow processing predicted students' mathematics performance at the adolescence stage. Campos, Almeida, Ferreira, Martinez and Ramalho (2013) also found working memory to be a significant predictor of student mathematics achievement among a sample of third grade Portuguese students. Although the studies included students of varying ages and from varying geographical locations and nationalities, they all yield consistent results.

Problem-solving ability, as a specific cognitive domain, is a significant predictor of student mathematics achievement (Primi et al., 2010; Vista, 2016; Wong & Ho, 2017). Vista (2016) investigated the role of problem-solving ability and reading comprehension skills in predicting growth trajectories of mathematics achievement in Australian students from third grade to eighth grade. Students' initial problem-solving ability predicted their initial level of mathematics achievement as well as the growth in mathematics achievement (Vista, 2016). The relationship between problem-solving ability and mathematics achievement was partially mediated by reading comprehension. The findings of this study support earlier findings by Geary (2011) and Primi et al. (2010) who found problem-solving to play a significant role in the learning of mathematics among seventh and ninth grade Portuguese students. Wong and Ho (2017) examined students' arithmetic word-problem solving among a sample of students from kindergarten to second grade in Hong Kong. The researchers' main aim was to identify correlates of student problem-solving component processes of worded mathematics problems. Student problem-solving abilities longitudinally predicted their computation and general mathematics achievement. However, domain-general skills predicted students' number-sentence construction whereas numerical-magnitude processing, word reading, and domain-general skills predicted arithmetic computation. Bjork and Bowyer-Crane (2013) investigated whether different cognitive skills underlie mathematical word problems and numeric operations. The study was conducted among a sample of second grade students in the United Kingdom. Bjork and Bowyer-Crane (2013) found reading comprehension and phonological awareness to be a significant predictors of students' mathematical word problem, while phonological awareness predicted students' performance on numerical operations. The studies have been conducted in varying countries, but have all shown consistent results. Problem-solving ability has been found to play a central role in mathematics achievement in general. However, different aspects of student problem-solving abilities have different effects on specific mathematics domain. An important general finding is that student problem-solving abilities are evident at an early age, and these abilities predict future mathematics achievement. These findings have important implications for designing mathematics instruction.

While cognitive abilities have been found to be associated with general mathematics achievement, other studies have found specific cognitive abilities to predict achievement in specific mathematics domain. Männamaa et al. (2012) examined the cognitive correlates of three domains of mathematics skills, namely: knowing, applying, and reasoning, or problem solving. Their aim was to identify potential deficits in cognitive areas that are associated with low mathematics achievement in specific domains among a sample of third grade students in the United States. Using confirmatory factor analyses (CFA), the researchers confirmed a four-factor model of mathematics skill: knowing-recalling, knowing-computing, applying and problem-solving. They found that verbal concepts contributed to the mathematics domains of knowing, applying and problem solving. In addition, verbal concepts and verbal reasoning were found to be most consistently associated with mathematics knowledge and problem-solving domains. Verbal working memory was also found to predict mathematics problem solving skills (Männamaa et al., 2012). Similar findings were reported by Zhang et al. (2017) who investigated the role of domain-general and numeric skills in predicting performance in arithmetic cognitive domains of knowing, applying and reasoning, among a sample of Finnish students. The researchers specifically examined the extent to which domain-general skills, such as spatial, language, rapid automatized naming (RAN) and memory at kindergarten and first grade predicted students' performance in fourth grade written computation, arithmetic word problems and arithmetic reasoning. CFA confirmed the four-factor model for the domain of general skills, spatial, language, rapid automatized naming and memory. These domain-general skills were found to play a central role in the development of students' arithmetic competence although they contributed independently to the learning of arithmetic. Domain-general skills played a mediating role between the development of numeric skills and arithmetic domain while spatial visualization was a unique predictor in arithmetic learning in the three arithmetic domains (Zhang et al., 2017). In both studies, the researchers conducted CFA to first confirm the factors, however, in the CFA conducted by Männamaa et al. (2012), the researchers were interested in confirming the mathematics skills, while Zhang et al. (2017) were interested in confirming the cognitive skills. However, both studies support the view that specific cognitive domains predict achievement in specific mathematics domain.

There is a lack of research on cognitive abilities and mathematics performance in the Caribbean. A search of the literature revealed that one such study was conducted at the primary level in Trinidad and Tobago. In a mixed method study, Khan (2017) investigated the proficiencies of students in the national Grade 4 mathematics examination in Trinidad and Tobago. The test included four content strands: number, measurement, geometry and statistics. These content strands were tested at three cognitive levels: recall, algorithmic thinking, and problem solving. Students' proficiencies were described according to four levels: below standard, nearly meets, meets, and exceeds. The data analysis included descriptive statistics and ANOVA repeated measures. Khan (2017) reported that the lower-performing group in the study demonstrated poor reading comprehension skills which affected their mathematics performance. Students performed poorly in the measurement strand, and in questions involving division and multiplication of algorithms. Overall, questions which required higher order thinking skills posed the greatest challenges for all students.

Summary and Conclusions

The literature on the mathematics content domains discussed above have established the importance of each strand in preparing students to become productive citizens and function effectively in society (Cass et al., 2003; Lake & Kemp, 2001; Pedersen, 2015; Reys et al., 2012; Schoenfeld, 1995; Serow et al., 2014; Tajudin & Chinnappan, 2016; Watson, 2001). Also of significance, is the integrative and synergic nature of the various content domains and the ways in which they contribute to a better understanding of each other and of mathematics in particular, and other curriculum learning areas in general (Goldstein, 2007; Van de Walle et al., 2013; Zhang et al., 2012). To emphasize the relatedness of the content domains, researchers consider measurement an amalgam of understanding numbers and geometry (Browning, Edson, Kimani, & Asian-Tutak, 2014). Several researchers and educators have advocated an integrative approach to the teaching of the various mathematics content domains for greater pedagogical benefits (Browning et al., 2014; Hurrell, 2015). Battista (2007) recommends incorporating students' experiences in the learning of geometry and measurement and engaging them in activities that allow them to explore and construct geometric ideas.

The relationship between cognitive abilities and mathematics achievement has been investigated using different aspects of cognitive abilities. While there have been some mixed results, most researchers have reported a strong positive relationship between these two variables across time, age, and country. Some general cognitive abilities, such as fluid reasoning and working memory were found to have a direct effect on student mathematical problem-solving skills. Bloom's taxonomy of educational objectives (Bloom et al., 1956) has provided a framework for the construction of mathematics assessment regionally and internationally, including TIMSS and CSEC mathematics. The cognitive domains adopted from Bloom's taxonomy are associated with general mathematics achievement as well as specific mathematics content domains and have been used to assess and report student mathematics achievement. Many large-scale mathematics assessments include mathematics content domains such as algebra, geometry, measurement, statistics and graphs, which are considered foundational concepts and are critical to day-to-day functioning and providing a foundation for higher learning.

While the relationship between cognitive abilities and mathematics achievement has been well established in the literature, most of the research were conducted among students in the United States, particularly at the primary level. From a Caribbean perspective, Khan (2017) has added to the literature with her research on cognitive abilities and mathematics achievement among students at the primary level in Trinidad and Tobago. Through this study, I hope to add to the literature on cognitive domains and mathematics achievement in specific content domains, thereby extending the investigation of the phenomena from a Caribbean perspective, particularly among secondary students in Saint Vincent and the Grenadines. It is hoped that the outcome of this study will provide educators with an in-depth understanding of students' cognitive by content achievement so that they can better target instructions to meet students' needs and abilities, with a view to improving mathematics achievement at the CSEC level. In

Chapter 3 I describe the method used to guide the study in seeking answers to the research question.

Chapter 3: Research Method

Introduction

The purpose of this study was to determine the extent to which the CSEC mathematics scores of high-scoring Vincentian students versus low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning differ across the content domains of algebra, geometry, measurement, statistics, and RFG. The study was quantitative in nature and used a nonexperimental, cross-sectional design. In this chapter, I describe the research design, rationale for the study, and variables investigated. The description includes the connection between the research design and the ways in which the design will advance knowledge in the discipline. I also give an explanation of the constraints consistent with the design. I outline the methodology of the study, which includes a description of the population, sample, sampling procedure, and operationalization of the variables. Following the operationalization of the variables, I describe the examination process, including the development and administration of the examination; marking, grading and reporting; and validity and reliability of the examination. I then outline the data analysis plan and the procedure for accessing the data, including the process of data cleaning. I discuss threats to internal validity, external validity, construct validity, and statistical conclusion validity, as well as ethical procedures, including treatment and protection of the data in accordance with the stipulations of the institutional review board (IRB). The chapter concludes with a summary of the design and methodology and a transition to Chapter 4.

Research Design and Rationale

In this nonexperimental quantitative study, I used a cross-sectional design based on the analyses of archival (secondary) data that comprised the scores of Vincentian students in the 2017 May/June CSEC mathematics examination. A cross-sectional study, also referred to a “snapshot” of a group of individuals, is an observational study in which the researcher simultaneously determines the exposure and outcome for each subject (Carlson & Morrison, 2009). The study design was cross-sectional in nature because data collection took place at one point in time and was used to compare two or more educational groups on a practice (Creswell, 2015; Ray, 2020). Advantages of the cross-sectional design include the ease of replication to other settings, and generating hypotheses, as well as the relatively low cost and the ability to study multiple outcomes from a single study (Bangdiwala, 2019). Constraints of the cross-sectional design include the inability to develop strong causal attributions, and the inability to establish change (Bono & McNamara, 2011; Spector, 2019). Although the cross-sectional design does not establish causal connection, knowledge of the association of the variables provides a basis for theory development and targeting intervention. Secondary data analysis is a methodological approach to data analysis in which the researcher uses data that are already in existence, such as a repository (Hosein, 2019). In the case of this study, the data existed in the CXC’s database. Advantages of using secondary data include the ability of the researcher to use a sample that spans a large geographical area and allows for the study of national trends unobtrusive to the study subjects. However, secondary data may not include all the variables of interest, or the data may not be captured in a form that is useful to the researcher (e.g., group level data versus individual data). Additionally, the data may be dated and may not reflect current trends. Also, secondary data do not allow for the establishment of causality (Bangdiwala, 2019; Bono & McNamara, 2011; Carlson & Morrison, 2009).

Variables

The study included four independent variables, which are categorical variables. Each independent variable had two sublevels. There were five dependent variables, measured at the continuous level.

Independent variables. The four variables were the three cognitive domains (knowledge, comprehension, and reasoning) and CoP (high-scoring students and low-scoring students).

Dependent variables. The five dependent variables were algebra scores, geometry scores, measurement scores, statistics scores, and RFG scores.

Methodology

Population

The population comprised 1,713 students from secondary schools and private institutions in Saint Vincent and the Grenadines who wrote the 2017 May/June CSEC mathematics examination. Mathematics at the CSEC level is compulsory in Saint Vincent and the Grenadines. The CSEC mathematics examination is usually written in Grade 11 (age 16 years) but may also be written by Grade 10 students who are more advanced. The examination may also be written by private candidates (students who are outside the regular secondary school setting and who may be attending private institutions, or students studying on their own without formal instruction).

Sampling and Sampling Procedures

I used stratified random sampling to select the sample for the study. Stratified random sampling is a method of sampling in which a population is divided into subgroups based on one or more variables central to the analysis of interest, then a random sample is drawn from each subgroup (Frankfort-Nachmias et al., 2015). In this study, I selected a stratified sample based on CoP: high-scoring students and low-scoring students in the 2017 May/June CSEC mathematics examination. I uploaded an Excel spread sheet file containing the scores of all the Vincentian students who wrote the 2017 May/June mathematics examination into the SPSS software. I ranked students by scores, from lowest to highest, 0 to 94. I then classified students with scores from 0 to 47 as low-scoring, and those with scores from 48 to 94 as high-scoring. The maximum available score was 94, and because 47 of 94 represents 50%, I used this as the criterion to separate the students into two groups. I categorized those students who scored to 50% as low-scoring, and those who scored above 50% as high-scoring. Of the 1,713 students who wrote the examination, 185 students met the criteria to be classified as high-scoring. I selected the high-scoring, then selected a random sample of 185 from the remaining 1,528 students classified as low-scoring.

I conducted G*Power analysis for two-way MANOVA with two levels and five dependent variables to determine an adequate sample size using alpha (α) of 0.05, a power of 0.80 and a small effect size ($f = 0.15$). Based on the assumptions, the G* Power analysis determined that a total sample size of 40 was sufficient (see Faul, Erdfelder, Buchner, & Lang, 2013). However, I used a total sample size of 370. I first selected the maximum number of high-scoring students available (185) and randomly selected a corresponding number of low-scoring students (185) to have equal numbers in each group. The measures taken in the sample selection were to ensure that all the sample requirements were met for a small effect size to control both the Type 1 error probability α and the Type 2 error probability $1-\beta$ (Mayr et al., 2007). The maximum score available on each cognitive domain (knowledge, comprehension, and reasoning) were 32, 34, and 28, respectively. I divided the scores on each cognitive domain into two strata as follows: knowledge, 0 to 16 (low-scoring) and 17 to 32 (high-scoring); comprehension, 0 to 17 (low-scoring) and 18 to 34 (high-scoring); reasoning, 0 to 14 (low-scoring) and 15 to 28 (high-scoring). I recoded the three independent variables, and the CoP, into categorical variables with two levels.

The data for the study were archival data from the CXC's database, comprising Vincentian students' scores in the 2017 May/June CSEC mathematics examination. The data comprised candidates' combined scores on Paper 01, the multiple-choice paper, and Paper 02, the essay paper, for questions that assessed the three cognitive domains in the five content domains of interest. The three cognitive domains are knowledge, comprehension, and reasoning, and the five content domains are algebra, geometry, measurement, statistics, and RFG. After the examination was written, the data became the property of the Ministry of Education, National Reconciliation and Information, Saint Vincent and the Grenadines, whereas the CXC was the custodian of the data. Written permission was requested and granted from both the owner and custodian of the data for use in the study (see Appendix F).

Walden University's IRB approval to advance to the data collection stage was received on April 4, 2020 (Approval No: 04-08-20-0402795; see Appendix G). After receiving approval, I sent a letter, via email, to the data manager in the Information System Department at the CXC, requesting the data. In the email, I attached the written permission received for the use of the data, from the CXC, and the Ministry of Education, National Reconciliation and Information, Saint Vincent and the Grenadines (see Appendix H). I followed up the email with a telephone conversation to ensure that the information was received, and the request was understood. I also sent the email to ascertain the earliest time by which I may receive the data. I received the data, via secure email on April 16, 2020. I immediately saved the data to my personal computer and backed up on two flash drives, which, when not in use, are password protected and kept in a locked filing cabinet.

Operationalization of Variables

There were four independent variables: the three cognitive domains (knowledge, comprehension, and reasoning), and CoP. Each independent variable had two levels. (high-scoring students and low-scoring students). The cognitive domain were the focal variables and CoP was the moderator variable. There were five dependent variables: algebra scores, geometry scores, measurement scores, statistics scores, and RFG scores. The operation of the independent variable of cognitive domain was demonstrated by students' ability to engage in various levels of cognitive processing as defined by

Bloom's taxonomy of learning objectives (Bloom et al., 1956) and operationalized in the CXC mathematics syllabus (Caribbean Secondary Examination Certificate, 2008). Knowledge represents foundational cognitive skill and requires students to recall facts, rules, definitions, and procedures, as well as perform simple computations. *Comprehension* is defined by the ability to engage in algorithmic thinking that involves translation from one mathematical mode to another, and the application of algorithms to solve familiar problems (CSEC, 2008). Reasoning is characterized by the ability to solve nonroutine problems, to make inferences and generalizations from given data and to analyze and synthesize information (CSEC, 2008). These cognitive domains provided the basis for the design of the CSEC mathematics examination, as well as for the analysis and reporting of students' results. For the CoP, a high-scoring student was denoted by a composite score from 48 to 94, and a low-scoring student was denoted by a composite score of 0 to 47 in the 2017 May/June CSEC mathematics examination. The dependent variables were the scores in the five mathematics content domains: algebra, geometry, measurement, statistics, and RFG. These content areas form part of the core of the CSEC mathematics curriculum. They are considered foundation concepts in mathematics that are required for everyday functioning in society. In this study, students' mathematics competence was determined by their ability to solve problems in these five content areas.

The Examination Process

The examination process entails the development and administration of the examination, marking and grading of the examination, reporting of the results, as well as the reliability and validity of the examination.

Examination development and administration. The 2017 May/June CXC mathematics examination was developed by a committee that comprised three mathematics content specialist and an assessment officer. The examination comprised two components: Paper 01 – a multiple choice paper, and Paper 02 – a constructed response paper. The multiple-choice paper was collated by the assessment officer, prior to the meeting, using pretested items from the item bank. The draft questions and accompanying key and mark schemes, or scoring rubrics for Paper 02, were written by the content specialist prior to the meeting, in accordance with a predetermined table of specifications (CXC, n.d.). A 5-day meeting was then convened to review and collate the draft examination papers. At the meeting, the questions were reviewed and the draft examination papers collated. The examination was moderated by an independent content expert and further reviewed and edited by three assessment officers and one copy editor (CXC, n.d.). The process of the examination development commenced 2 years prior to the administration of the examination. The mathematics examination was administered as a paper and pencil test to students in the 19 participating territories simultaneously (CXC, n.d.).

Examination script marking. The Paper 01 was machine scored and the Paper 02 was marked on screen by mathematics teachers in the various Caribbean countries. The marking of the examination scripts began with the process of standardization, where all potential markers were oriented to the scoring rubric (CXC, n.d.). The committee that prepared the examination, as well as other mathematics content specialist, engaged in a process of standardization during a 4–day period. The process was guided by the assessment officer. Standardization of Paper 02 involved reviewing the examination papers, key and mark scheme, or scoring rubric to ensure that the scoring rubric accurately reflects the tasks required by the examination questions and amending the scoring rubric where necessary (CXC, n.d.). Using the electronic marking tool, the committee selected a random sample of students' responses from traditional high-scoring schools, average scoring schools, and low-scoring schools in various territories. The committee then used the scoring rubric to mark the responses independently, then compared and discussed the scores to arrive at final agreed (definitive) scores. Alternative, valid methods used by students to solve the problems were accepted and incorporated into the scoring rubric (CXC, n.d.). After the committee completed standardization, they in turn standardized a group of experienced teachers, referred to as seed makers. The seed makers assisted the committee in marking additional responses, which were reviewed by the committee and classified as seeds. The seeds were used as quality standards, to judge the accuracy of the markers marking (CXC, n.d.). The committee also selected a set of responses that were used by the markers as practice responses, standardization responses, and an additional standardization (STM) responses, for markers who had to be restandardized. Prior to the markers being standardized, they were required to attend a virtual standardization meeting to discuss the marking of the responses to which they were assigned to mark. During the meeting, their supervisor oriented the markers to the scoring rubric, highlighting any nuances and peculiarities in the scoring rubric, for example, the award of partial credit for partially correct responses (CXC, n.d.).

The markers began the process by using the scoring rubric to mark the assigned practice responses. After marking each practice response, the markers were allowed to see the model responses that were marked by the committee. The markers then compared their marked responses against the model responses so that they could determine the level of accuracy and the areas that needed improving (CXC, n.d.). After the markers completed the practice marking, they then engaged in standardization. For the standardization, the markers were required to mark at least eight out of 10 responses within an agreed range (tolerance level) from the model responses. The markers who achieved this objective were automatically approved to engage in live marking. Markers who did not achieve the objective received feedback from their supervisors and were required to restandardize using a different set of responses (CXC, n.d.). Markers who failed standardization twice were not allowed to engage in the marking exercise. Quality assurance during marking included supervisors reviewing and or remarking responses marked by markers. Additionally, a set of responses referred to as seeds, which were responses previously marked by the supervisors and approved by the committee, appeared at random to the markers. If markers marked two consecutive seeds out of tolerance, they were suspended from the marking exercise (CXC, n.d.).

Grading and reporting of examination results. After the marking of all responses was completed, the scores were analyzed by the assessment officer and the examining committee, and grades awarded based on predetermined criteria for the award of grades (CXC, n.d.). These scores and grades were reviewed by a technical advisory committee who interrogated the examination process, commencing with the development of the examination through standardization, marking and the awarding of grades. The final sanctioning of the scores and grades was done by the final awards committee comprising representatives from the various territories, headed by the chairman of the organization (CXC, n.d.).

Reliability and validity of the CSEC mathematics examination. Reliability and validity are necessary features of educational assessment required for making decisions regarding learners' ability. Assessment validity refers to "the degree to which test-based inferences about students are accurate" (Popham, 2000, p. 94). The more evidence of validity, the more confidence one can place on the score-based inferences. There are three essential kinds of evidence that determine whether the inferences one makes from an educational assessment procedure are valid. The three kinds of evidence of validity are: content related evidence of validity, criterion related evidence of validity, and construct related evidence of validity. Content related evidence of validity refers to the extent to which an assessment procedure adequately represents the content of the assessment domain being sampled" (Popham, 2002, p. 52). Criterion related evidence of validity relates to whether performance on one assessment procedure accurately predicts performance on an external criterion, while construct related evidence of validity relates to whether there is empirical evidence that an inferred construct exists and a given assessment procedure is measuring the inferred construct accurately (Popham, 2002).

Content related evidence of validity is established by using a test blueprint or table of specifications. The table of specifications specifies the cognitive process and the content to be covered by the test (Thorndike, 1997). The development of the 2017

May/June CSEC mathematics examination was based on the CSEC (2008) mathematics syllabus which was developed by the CXC and taught in the 19 territories. The examination process was guided by specimen papers and a table of specifications which together form a blueprint for the examination. The use of the table of specifications helps to establish validity of the examination by ensuring that the syllabus objectives and content are proportionately represented as stipulated by the syllabus.

Nitko (2004) defined reliability as the degree to which students' results remain consistent over replication of an assessment procedure. Reliability is established when students' assessment results are the same in any of three situations; when they complete the same task(s) on two or more different occasions, two or more teachers mark their performance on the same task(s), or when they complete two or more different but equivalent tasks on the same or different occasions (Nitko, 2004). The reliability of the mathematics examination was established based on Nitko's (2004) third condition as the students completed two components of the examination on two different occasions during the month of June 2017. The two components were the multiple-choice paper and the constructed response paper, both papers included the same content areas and were tested at the same cognitive levels. The CXC used the Kuder Richardson Formula 20 to estimate the reliability of the multiple-choice paper and Cronbach's alpha to estimate the reliability of the constructed response paper and the whole examination. The reliability estimates for the three components were .91, .93, and .95 respectively (CXC, 2018). This means that students' performance scores were consistent across the two components of the test. Hence, the test can be considered to have produced reliable scores. While reliability does not guarantee validity, it is a necessary condition for validity (Nitko, 2004; Popham, 2002). Based on the reliability estimates, the achievement domain was consistently measured. The table of specification ensures that the content domain was adequately sampled, based on the assessment requirements stipulated in the syllabus (CSEC, 2008), so that accurate score-based inferences can be made about students' mathematics abilities.

Data Analysis Plan

I used secondary data in this study. Secondary data analysis is a methodological approach to data analysis in which a researcher uses data that are already in existence, such as a repository (Hosein, 2019). Some of the advantages of a secondary data analysis approach include access to larger datasets than would be otherwise feasible given the usual constraints of time and cost. Also, an existing dataset allows the data to be used parsimoniously (Hosein, 2019). However, the data are not always captured in a form required to answer the research question and sometimes proxies must be used (Hosein, 2019). The secondary data for this study were retrieved from the CXC's database and comprised the scores obtained by Vincentian students in the 2017 May/June CSEC mathematics examination. SPSS version 25 was used to analyze the data to answer the research question.

The 2-way MANOVA has 10 assumptions that must be considered when choosing this statistical analysis. The first three assumptions relate to the study design and should be met prior to conducting the study (Ates, Kaymanz, Kale, & Tekindal, 2019; Ito, 1980; Pituch & Stevens, 2016). The other seven assumptions relate to how the data fits the 2-way MANOVA model and can be tested using SPSS. The assumptions are as follows:

- Assumption 1. There should be two or more dependent variables that are measured at the continuous level.
- Assumption 2. There should be two or more independent variables where each independent variable consists of two or more categorical independent groups.
- Assumption 3. There should be independence of observation. That is, there should be no relationship between the observation in each group of the independent variable or between the groups themselves.
- Assumption 4. There should be a linear relationship between the dependent variables for each group of independent variables.
- Assumption 5. There should be no multicollinearity. That is, the dependent variables should be moderately correlated with each other.
- Assumption 6. There should be no univariate and multivariate outliers. Univariate outliers are values of the dependent variable that are unusual within each group of the independent variable, whereas multivariate outliers are cases that have unusual combination of scores on the dependent variables.
- Assumption 7. There should be multivariate normality. That is, normally distributed data for each combination of the independent variables for all dependent variables.

- Assumption 8. There should be an adequate sample size. That is, each cell of the matrix should have at least as many cases as there are dependent variables.
- Assumption 9. There should be homogeneity of variance-covariance matrices. That is, variances and covariances of the dependent variable in each cell of the design (i.e., group combination) should be equal in the population.
- Assumption 10. There should be homogeneity of variances. That is, there should be equal variances in each cell of the design for each dependent variable (Nimon, 2012; Pituch & Stevens, 2016).

Data cleaning helps to improve data normality, linearity, and homoscedasticity (Osborne, 2010; Sakia, 1992). Data cleansing involves removing out-of-range numbers that can skew the results (Chan, 2003). To reduce the chance of committing either a Type I or Type II error, I performed data cleansing in SPSS. Data cleansing included generating frequency tables and inspecting each for out-of-range values and generating descriptive statistics using skewness and kurtosis; generating histograms and percentage plots (P-P plots) and performing inferential test of normality such as Kolmorogov-Smirnow and Shapiro-Wilk's W test. According to Chan (2003), these tests help to improve data normality and result in the production of more appropriate descriptive statistics and the application of correct statistical tests, and thereby improving the results of the analyses.

Research Question: How do the CSEC mathematics scores of high-scoring Vincentian students and low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning differ across the content domains of algebra, geometry, measurement, statistics, and RFG?

H_0: There are no differences in the CSEC mathematics scores between high-scoring Vincentian students and low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning across the content domains of algebra, geometry, measurement, statistics, and RFG.

*H*a: There are differences in the CSEC mathematics scores between high-scoring Vincentian students and low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning across the content domains of algebra, geometry, measurement, statistics, and RFG.

Analysis of data included descriptive statistics, 2-way MANOVA statistical analysis for differences between groups, and follow-up 2-way ANOVA. The results were interpreted at the .05 α level of significance. The 2-way MANOVA is a statistical test used to test the interaction effect between two independent variables and two or more combined dependent variables. Hence, the 2-way MANOVA was used to test the null hypothesis (H_0.) – there are no differences in the CSEC mathematics scores between high-scoring Vincentian students and low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning across the content domains of algebra, geometry, measurement, statistics and RFG.

Threats to Validity

Threats to validity refers to statistical and design issues that threaten the research and could cause the researcher to draw false conclusion from the data (Creswell, 2015). Cook and Campbell (1979) identified four facets of research validation. These include internal validity, external validity, construct validity, and statistical conclusion validity. In any type of research design, it is paramount that researchers account for threats to all forms of validity to have meaningful research results. In the cross-sectional cohort design, the researcher conducts a cross-sectional sampling to obtain a study cohort and then performs a retrospective assessment of the history of exposure and outcomes in the members of that cohort (Hudson, Pope, Jr, & Glynn, 2005).

Threats to external validity are problems that threaten the generalizability of the findings of one study to other setting, persons, and situations (Frankfort-Nachmias et al., 2015). Threats to external validity include interaction of selection and treatment, interaction of setting and treatment, and interaction of history and treatment (Creswell, 2015). Interaction of selection involves the inability of a researcher to generalize the findings of a study beyond the group that is studied (Creswell, 2015). To reduce this threat, I used stratified random sampling to ensure that sub-groups of high-scoring students and low-scoring students in the sample of Vincentian students represent the subgroups of high-scoring students and low-scoring students in the Vincentian student population. Interaction of setting and treatment arises from the inability of a researcher to generalize from the study setting to other settings (Creswell, 2015). The data for the study included students who attended regular secondary schools, students who attended private institutions that are not classified as schools, and students who did not attend any formal institutions and wrote the examination as private candidates. To reduce the threat of interaction of setting and treatment, I did not generalize the findings to any one group of subjects. That is, I generalized the findings of the study to the total sample studied and not to any subgroups such as private candidates or students in school. Interaction of history and treatment develops when the researcher tries to generalize the findings of one study to past or future situations (Creswell, 2015). The data for this study comprised the scores of Vincentian students who wrote the 2017 May/June CSEC mathematics examination. To reduce the threat of interaction of history and treatment, I generalized the findings of the study only to the cohort of students who wrote the examination in May/June 2017 and no cohort who wrote the examination in any other sitting.

Threats to internal validity are issues or problems with procedures or participants that can compromise the inferences that are drawn from the study. Threats to internal validity in the cross-sectional design include measurement errors, bias, chance, and non-ignorable exiting and inflation of causal inference due to common method variance (CMV) (Hartung & Touchette, 2009; Jackson et al., 2005; Jackson, O'Callaghan, & Adserias, 2014). Measurement errors are internal threats to the validity of a study if unaccounted for in the analysis, could result in spurious findings, or CMV and erroneous casual inference (Jackson et al., 2014). CMV are variance attributable to the methods used to measure the construct rather than to the construct being measured. These measurement methods may include using a single rater, item characteristics, item context, and measurement context (Campbell & Fiske, 1959; Podsakoff et al., 2003). As a potential source of measurement error, CMV may inflate or deflate the correlation among research variables, thereby threatening the validity of the conclusions drawn about the relationships between the measures of different constructs (Reio, 2010). CMV in quantitative studies can be controlled by strengthening the procedural design of the study, and by using statistical controls (Podsakoff et al., 2003; Rindfleisch et al., 2008). In this study, I minimized measurement errors by using the same measurement instrument, in the form of a mathematics examination. All students wrote the examination at the same time, and teachers used the same scoring rubric to mark the students' scripts. To ensure that the markers applied the scoring rubric consistently, a supervisor remarked a sample of scripts markers by each marker. The data manager used the method to retrieve all the students' scores from the CXC's database.

Selection bias occurs when subjects are selected for a study in such a way that creates false association, whereas information bias occurs when the method of data collection between groups is significantly different (Hartung & Touchette, 2009). Selection bias creates a systematic error in the measurement of the variables. Chance as a threat to validity, occurs when random variations result in observable differences (Hartung & Touchette, 2009). Non-ignorable exiting refers to situations where subjects exit the study before the time of evaluation of the outcome (Hudson et al., 2005). Non-ignorable exiting does not apply to this study because the data were collected at one point, hence, there were no threats to participants exiting the study. Thus, reducing threats to internal validity. I minimized the likelihood of selection bias and chance bias by increasing the sample size beyond the recommended size determined by the

G*statistics. I used the G*Power analysis to determine an adequate sample size for the study, which was determined to be 40, however, I used a sample size of 370. This larger sample size ensured that the all the sample requirements were met for a small effect size and controlled both the Type I error probability α and the Type 2 error probability $1-\beta$ (Mayr et al., 2007).

Threats to construct validity are problems relating to the independent variable and the dependent variable used in the study that threaten the ability of the researcher to make correct inferences (Creswell, 2015). According to García-Pérez (2012), construct validation is established by using well-established definitions and measurement procedures for variables. To reduce threats to construct validity, I have provided operational definitions of the independent and dependent variables used in the study, and outlined clearly how the variables were measured. Since the study used secondary data, the measure of the variables was established prior to conducting the study.

Statistical conclusion validity (SCV), pertains to the extent to which statistical analyses of the data of a research study can reasonably reveal a link (or lack thereof) between the independent and dependent variables (Cook & Campbell, 1979). SCV was summarized as “inferences about whether it is reasonable to presume covariation given a specified α level and the obtained variances” (Cook & Campbell, 1979, p. 41). Given this definition, SCV was seen as including three aspects: the statistical power of the study to detect an effect, the risk associated with revealing an effect that does not exist, and the ability to confidently estimate the magnitude of the effect. SCV is also concerned with sources of random error and the appropriate use of statistics and statistical tests (Cook &

Campbell, 1979). However, Cook and Campbell (1979) acknowledged that the potential occurrence of Type I and Type II errors cannot be prevented as they are an essential and inescapable consequence of the statistical decision theory underlying significance test. Further, these errors only affect SCV when there is a meaningful difference between the assumed and actual probability (García-Pérez, 2012). There are two main threats to SCV. The first threat occurs when the design used to collect the data does not match the characteristics of the data analysis and the statistical analyses applied in analyzing the data are methodically inadequate and cannot logically provide an answer to the research questions. The second threat occurs when the appropriate statistical tests are applied in analyzing the data, but the tests violate the stated risks probabilities (García-Pérez, 2012). To reduce these threats, I ensured that the three assumptions of the 2-way MANOVA which relate to the design of the study, were met prior to conducting the study. These assumptions are there must be two or more dependent variables that are measured at continuous levels; there must be two or more independent variables consisting of two or more categorical, independent groups; and there must be independence of observation (Pituch & Stevens, 2016). The other seven assumptions were tested prior to conducting the data analysis.

Ethical Procedures

I conducted the research in accordance with Walden University IRB requirements. I received written permission from the Ministry of Education, National Reconciliation and Information, Saint Vincent and the Grenadines, the owners of the data, and from the CXC, the custodian of the data, for the use of the mathematics scores of Vincentian students who wrote the CSEC mathematics examination in May/June 2017. Initial permission was also received from the CXC in 2018 for use of descriptive statistics and CXC related materials to establish the problem. I followed all necessary protocols in receiving and handling the data. In the letters to the Ministry of Education, National Reconciliation and Information, Saint Vincent and the Grenadines and the CXC, I indicated the purpose of the study, the type of data required and how I will use the data. I also share these specifications with the data manager, when I requested the data. Hence, the data manager cleaned the data by removing all student identification, including student age, registration number, and school. The data manager sent the data via the CXC's secure email and I immediately saved the data to my personal computer. I also saved backed up copies on two flash drives that I password protected and keep in a locked filing cabinet. I will retain the data 5 years and then destroyed, as per Walden University IRB requirement. While I acquired the data for the study from the organization where I work, I did not have any contact with the construction, administration, marking, or grading of the 2017 May/June CSEC mathematics examination, neither do I know the students personally, hence there is no basis for researcher bias. Additionally, I did not have any direct contact with the students' mathematics scores. The data manager collated the data which consisted of students' mathematics scores for algebra, geometry, measurement, statistics, and RFG, by cognitive domains (knowledge, comprehension, and reasoning). Construct validity refers to how well a researcher operationalizes a construct. That is, whether the researcher accurately transforms or translates concepts, ideas and behavior into functioning and operating reality (Trochim, 2006). I operationalized the variables in this study based on the way they are defined in the CSEC mathematics syllabus and used in the examination.

Summary

In this quantitative study, I used a nonexperimental, cross-sectional design to determine the extent to which the CSEC mathematics scores of high-scoring Vincentian students versus low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, reasoning differ vary across the content areas of algebra, geometry, measurement, statistics and RFG. The sample comprised 185 students classified as high-scoring and 185 students classified as low-scoring in the 2017 May/June CSEC mathematics examination based on their composite score in the five content domains. The study included four independent variables, each with two sub-levels, and five dependent variables. The independent variables were: knowledge, comprehension, reasoning, and CoP. The sub-levels were levels of scoring (high, low). The dependent variables were students' scores in the five content domains (algebra, geometry, measurement, statistics, and RFG). Data analysis included descriptive statistics, 2-way MANOVA and follow up 2-way ANOVA. The results were reported at the .05 α level of significance, consistent with most educational research. I conducted the study in accordance with Walden University's IRB guidelines by seeking permission to use data comprising students' examination scores, and adhering to the guidelines regarding the treatment and confidentiality of the data. Threats to the validity of this quantitative cross-sectional study included measurement errors, construct validity, and SCV, which I minimized by the procedures I utilized in the study. I ensured that ethical standards were met by adhering to the Walden University IRB’s requirements regarding the collection and treatment of data, as well as reducing researcher bias. The next chapter, Chapter 4 addresses the data analysis and presentation.

Chapter 4: Results

Introduction

The purpose of this quantitative, nonexperimental study was to determine the extent to which the CSEC mathematics scores of high-scoring Vincentian students and low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning differ across five content domains of algebra, geometry, measurement, statistics, and RFG. I used archival data that comprised the scores of Vincentian students in the 2017 May/June CSEC mathematics examination. This chapter includes a review of the research question and hypotheses and a description of the data collection process, including the population, data cleansing, and selection and composition of the sample. These sections are followed by the statistical analyses and findings of the study. The chapter concludes with a summary of the outcome of the analyses.

Research Question and Hypotheses

Research Question: How do the CSEC mathematics scores of high-scoring Vincentian students and low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning differ across the content domains of algebra, geometry, measurement, statistics, and RFG?

*H*o: There are no significant differences between the CSEC mathematics scores of high-scoring Vincentian students and low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning across the content domains of algebra, geometry, measurement, statistics, and RFG.

*H*a: There are significant differences between the CSEC mathematics scores of high-scoring Vincentian students and low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning across the content domains of algebra, geometry, measurement, statistics, and RFG.

Data Collection

On February 14, 2020, I wrote official letters to the Ministry of Education, National Reconciliation and Information, Saint Vincent and the Grenadines (the owner of the student data), and the CXC (the custodian of the data) seeking permission to use Vincentian students' mathematics scores in the 2017 May/June CSEC mathematics examination to conduct research (see Appendices C and E). Both letters outlined the purpose of the research, the type of data required, and how the data would be treated to ensure anonymity of the students. I received written permission from Saint Vincent and the Grenadines and the CXC on March 18 and 20, 2020, respectively (see Appendices D and F). These letters of permission were submitted as part of the application to the Walden University IRB on March 27. I received IRB approval on April 4 (Approval No: 04-08-20-0402795) to advance to the data collection stage (see Appendix G). On April 9, 2020, I wrote an email to the data manager at the CXC requesting the data. I attached the letters of approval for data use from Saint Vincent and the Grenadines and the CXC, including a table indicating how the data should be organized (see Appendix H). I received the data in an excel spreadsheet via the CXC's email on April 16, 2020. I saved the data on my personal computer and saved backup files on two flash drives, which I kept in a locked in a filing cabinet. I password protected all files.

In my data plan, I indicated that I would receive the Excel file with the data on a flash drive. However, by the time of data collection, Barbados had declared a state of emergency as a result of the COVID-19 pandemic, and employees of the CXC were working remotely. As a result, the safest and most efficient way of receiving the data was via the CXC's secure email. The data contained students' scores on each cognitive domain (knowledge, comprehension, reasoning) for each content domain (algebra, geometry, measurement, statistics, and RFG), as well as the composite score. The data manager had performed data cleansing prior to releasing the data. The data file did not contain any student identification, such as student number or school; the cases were numbered from 1 to 1,713. In addition, there were no missing data; all students had scores for all components of the examination, including scores for each cognitive domain and content domain. Any student who did not complete both components (Paper 01 and Paper 02) were removed from the population. In addition, there were no significant outliers; therefore, no additional data cleansing was required.

I used G*Power analysis to determine an appropriate sample size for the study. The G*Power analysis with alpha of 0.05, power of 0.80, and small effect size ($f = 0.15$) indicated the minimum sample size was 40. As I had access to a population of 1,713 students, I proposed using a sample size of 400, including 200 in each of the CoP (high and low performance) groups. The reason for choosing this sample size was to ensure that all the sample requirements would be met for a small effect size to control for both the Type 1 and Type 2 errors (see Mayr et al., 2007). Based on the range of scores for each group, only 185 students were classified as high CoP. To maintain an equal number of students in each group, I selected a random sample of 185 from the population of 1,528 students classified as low CoP. Therefore, the total sample size was 370. I divided the scores on each of the four independent variables (knowledge, comprehension, reasoning, and CoP) into two strata (high and low) and coded them as categorical variables with two levels. Table 6 shows the range of scores used to determine the high and low groups for each of the independent variables. Table 7 shows the composition of the groups for each of the independent variables.

Table 6

Range of Scores for High and Low Groups for the Independent Variables

Independent variable	Low	High
Knowledge	0-16	17-32
Comprehension	0-17	18-34
Reasoning	0-14	15-28
CoP	0-47	48-94

Note. The score in the upper range of the high group represent the maximum score attainable for each variable.

Table 7

High and Low Groups for the Independent Variables

Independent variable	Level	Number of students	Percentage of students
Knowledge	Low	221	59.7
	High	149	40.3
Comprehension	Low	174	47.0
	High	196	53.0
Reasoning	Low	264	71.4
	High	106	28.6
CoP	Low	185	50.0
	High	185	50.0

Results

Descriptive Statistics

Descriptive statistics were generated to determine whether there were missing data or outliers. Table 8 shows the means and standard deviations, as well as the range of scores achieved by students on each content domain, and the maximum available score for each content domain.

Table 8

Descriptive Statistics for the Mathematics Content Domains

Content domain	Mean	St. dev	Range	Maximum
Algebra	10.45	5.01	1-20	20
Geometry	8.67	4.67	1-20	20
Measurement	8.04	5.16	0-20	20
Statistics	7.45	3.80	0-16	16
RFG	7.35	4.69	0-18	18

Note. The maximum represents the maximum available mark for each content domain.

Assumptions Testing

The two-way MANOVA has 10 assumptions that must be considered (Nimon, 2012; Pituch & Stevens, 2016). The first three assumptions relate to the design of the study and were met prior to conducting the study.

Assumption 1: Two or more dependent variables. There should be two or more variables that are measured at the continuous level. In this study there were five dependent variables measured at the continuous level. These dependent variables were students' mathematics scores in the 2017 May/June CSEC mathematics examination and comprised scores in algebra, geometry, measurement, statistics, and RFG.

Assumption 2: Categorical independent variables. There should be two or more independent variables where each independent variable consists of two or more categorical independent groups. In this study, there were four independent variables:

knowledge, comprehension, reasoning, and CoP. Each variable comprised two sublevels (high, low).

Assumption 3: Independence of observation. There should be independence of observation. That is, there should be no relationship between the observation in each group of the independent variable or between the groups themselves. The students in this study were categorized in one of the two sublevels for each independent variable. Students were in either the high group or the low group, for each independent variable, but not both.

The remaining seven assumptions were tested prior to conducting the two-way MANOVA. Assumptions testing was conducted for linearity, multicollinearity, univariate outlier and multivariate outlier, normality, adequate sample size, homogeneity of variance-covariance matrices, and homogeneity of variances.

Assumption 4: Linearity. There should be a linear relationship between the dependent variables for each group of independent variables. Using paired combinations of the independent variables, I tested for linearity to determine whether there were univariate outliers for each combination of the four independent variables. Forty scatterplots were generated. Inspection of the scatterplots indicated a linear relationship between 34 of the 40 pairs of dependent variables across each level of the independent variables. Hence, the assumption of linearity was met. Figure 3 shows the scatterplot for algebra and geometry for levels of knowledge. See Appendix I for additional samples of the scatterplots.

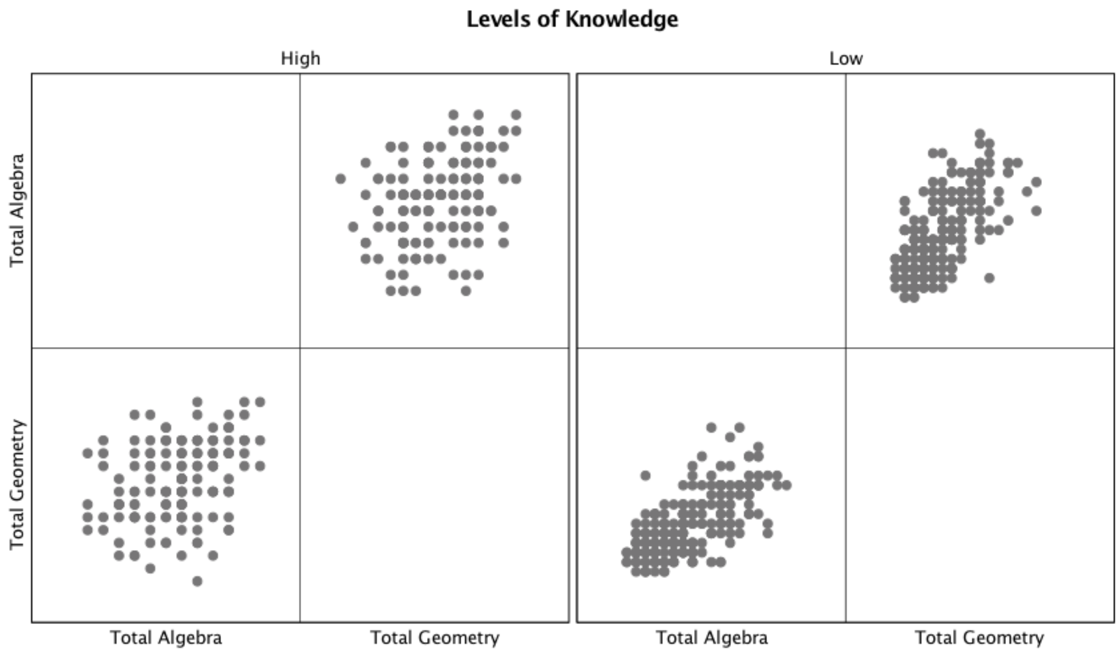


Figure 3. Scatterplots for algebra and geometry for levels of knowledge.

Assumption 5: Multicollinearity. There should be no multicollinearity. That is, the dependent variables should be moderately correlated with each other. To satisfy the assumption of multicollinearity, the dependent variables must be related but not highly correlated (Pituch & Stevens, 2016). Pearson correlation was used to test for multicollinearity. The results of the correlation (|r| <0.09) indicated that the variables were positively related; the strength of the correlation ranged from 0.23 to 0.79. Therefore, there was no evidence of multicollinearity, and this assumption was met. Table 9 shows the results of the correlation.

Table 9

Pearson Correlation for the Content Domains

Content domain		Algebra	Geometry	Measurement	Statistics	RFG
Algebra	Pearson correlation	1	.793**	.735**	.702**	.831**
	sig. (2-tailed)		.000	.000	.000	.000
	N	370	370	370	370	370
Geometry	Pearson correlation	.793**	1	.723**	.679**	.786**
	sig. (2-tailed)	.000		.000	.000	.000
	N	370	370	370	370	370
Measurement	Pearson correlation	.735**	.723**	1	.582**	.743**
	sig. (2-tailed)	.000	.000		.000	.000
	N	370	370	370	370	370
Statistics	Pearson correlation	.702**	.679**	.582**	1	.688**
	sig. (2-tailed)	.000	.000	.000		.000
	N	370	370	370	370	370
RFG	Pearson correlation	.831**	.786**	.743**	.688**	1
	sig. (2-tailed)	.000	.000	.000	.000	
	N	370	370	370	370	370

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

Assumption 6: Univariate and multivariate outliers. There should be no univariate and multivariate outliers. There should be no univariate outliers in each group combination of the independent variable (i.e., for each cell of the design) for any of the dependent variables. Multivariate outliers are cases that have unusual combination of scores on the dependent variables. The 2-way MANOVA is sensitive to both univariate outliers and multivariate outliers (Tabachnick & Fidell, 2014). Univariate outliers are values of a dependent variable that are unusual within each group of the independent variables, whereas multivariate outliers are data points that have unusual combination of values within the dependent variables (Pituch & Stevens, 2016) To test for univariate outliers, I generated boxplots for each combination of the dependent variables and independent variables. Inspection of the boxplots indicated the presence of univariate outliers. Therefore, the assumption of no univariate outlier was not met. Figure 4 shows a boxplot for algebra scores. See Appendix J for other samples of boxplots. The Mahalanobis distance was calculated as part of the linear regression analysis to test for multivariate outliers. There were no multivariate outliers in the data as assessed by Mahalanobis distance ($p > .001$). The Mahalanobis recorded for the data was 19.37, which was less than the critical value of 20.52. Therefore, the assumption of no multivariate outliers was met.

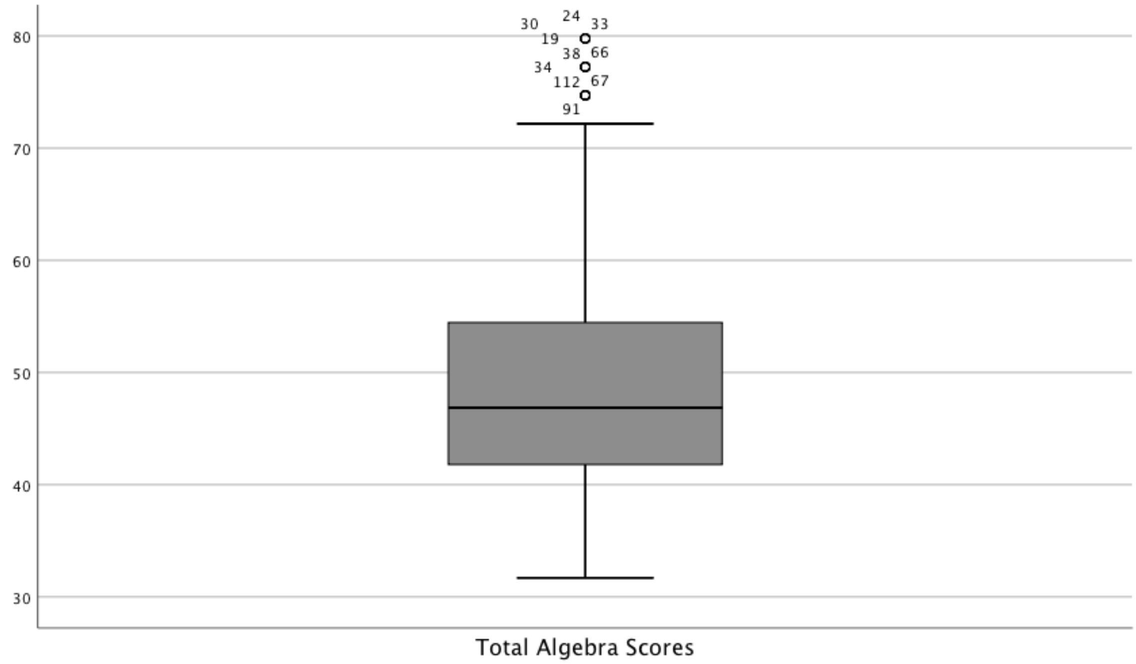


Figure 4. Boxplot for algebra scores.

Assumption 7-Multivariate normality. There must be multivariate normality. That is, normally distributed data for each combination of the independent variables for all dependent variables. According to the Shapiro-Wilk test for normality with a sample of 370 students, the assumption of normality was not met because the scores on the five cognitive domains (algebra, geometry, measurement, statistics, and RFG) were considered to have a non-normal distribution across knowledge, comprehension, reasoning, and CoP (p <.05). The non-normal distribution was also confirmed through the inspection of histograms. The results of the Shapiro-Wilk test for normality are presented in Tables 10, 11, 12, and 13. The distribution of scores represented the actual performance of students, and Pituch and Stevens (2016) posit that in cases where this assumption is violated, but the value for kurtosis is positive, departure from normality is not expected to have much effect on power and should not be cause for concern. Hence, I did not consider the violation of this assumption to be a threat to the analysis, and I proceeded with the analysis of the data.

Table 10

Tests of Normality: Levels of Knowledge and Content Domains

Content domain	Level of knowledge	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Algebra	Low	.118	221	.000	.953	221	.000
	High	.116	149	.000	.974	149	.006
Geometry	Low	.142	221	.000	.935	221	.000
	High	.120	149	.000	.973	149	.005
Measurement	Low	.175	221	.000	.903	221	.000
	High	.098	149	.001	.970	149	.003
Statistics	Low	.116	221	.000	.958	221	.000
	High	.125	149	.000	.961	149	.000
RFG	Low	.151	221	.000	.927	221	.000
	High	.103	149	.001	.979	149	.022

Note. a. Lilliefors Significance Correction.

Table 11

Tests of Normality: Levels of Comprehension and Content Domains

Content domain	Level of comprehension	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Algebra	Low	.107	174	.000	.961	174	.000
	High	.112	196	.000	.978	196	.003
Geometry	Low	.153	174	.000	.943	174	.000
	High	.101	196	.000	.977	196	.002
Measurement	Low	.151	174	.000	.914	174	.000
	High	.087	196	.001	.973	196	.001
Statistics	Low	.130	174	.000	.951	174	.000
	High	.108	196	.000	.970	196	.000
RFG	Low	.167	174	.000	.910	174	.000
	High	.110	196	.000	.982	196	.011

Note. a. Lilliefors Significance Correction.

Table 12

Test of Normality: Levels of Reasoning and Content Domains

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
Content domain	Level of reasoning	Statistic	df	Sig.	Statistic	df	Sig.
Algebra	Low	.106	264	.000	.957	264	.000
	High	.107	106	.005	.970	106	.015
Geometry	Low	.131	264	.000	.946	264	.000
	High	.152	106	.000	.965	106	.007
Measurement	Low	.154	264	.000	.934	264	.000
	High	.080	106	.094	.970	106	.017
Statistics	Low	.106	264	.000	.961	264	.000
	High	.158	106	.000	.929	106	.000
RFG	Low	.139	264	.000	.936	264	.000
	High	.117	106	.001	.969	106	.013

Note. a. Lilliefors Significance Correction.

Table 13

Tests of Normality: Category of Performance and Cognitive Domains

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
Content domain	CoP	Statistic	df	Sig.	Statistic	df	Sig.
Algebra	LowCoP	.104	185	.000	.964	185	.000
	HighCoP	.105	185	.000	.980	185	.010
Geometry	LowCoP	.151	185	.000	.946	185	.000
	HighCoP	.105	185	.000	.975	185	.002
Measurement	LowCoP	.165	185	.000	.886	185	.000
	HighCoP	.090	185	.001	.971	185	.001
Statistics	LowCoP	.124	185	.000	.954	185	.000
	HighCoP	.104	185	.000	.969	185	.000
RFG	LowCoP	.159	185	.000	.921	185	.000
	HighCoP	.119	185	.000	.978	185	.006

Note. a. Lilliefors Significance Correction.

Assumption 8: Adequate sample size. There should be an adequate sample size. In conducting a 2-way MANOVA, each cell of the design must have at least as many cases as there are dependent variables (Pituch & Stevens, 2016). To test this assumption, I generated descriptive statistics. This assumption was met except for the combination of CoP and high levels of reasoning, and high levels of knowledge and low CoP (see Appendix K). All other cells had more than the minimum five students required. The small numbers that are noted in some cells are expected because the students who scored high on knowledge and reasoning naturally scored high on the examination overall. It was therefore unlikely for a student to score high on a content domain and be classified as low CoP. Hence, I proceeded with the analyses.

Assumption 9: Homogeneity of variance-covariance matrices. There should be homogeneity of variance-covariance matrices. That is, variances and covariances of the dependent variable in each cell of the design (i.e., group combination) should be equal in the population. I used Box's M test of equality of covariances to determine whether the variances and covariances of the dependent variables (algebra, geometry, measurement, statistics, and RFG) for each combination of the cognitive domains (knowledge, comprehension, and reasoning) and CoP are equal in the population. This assumption was met for three of the five dependent variables (measurement, statistics, and RFG) as assessed by Box's M test ($p < .001$). The assumption of homogeneity of variance-covariance matrices is very restrictive and it is unlikely that this assumption would be satisfied in practice (Konietschke, Bathke, Harrar, & Pauly, 2015; Pituch & Stevens, 2016). Bathke et al. (2018) added that plausible violation of this assumption that may occur in practice may not have much of an effect on power. Hence, I did not consider the partial violation of this assumption to be a threat to the analysis, and I proceeded with the analysis. The results of the test are indicated in Table 14.

Table 14

Box's Test of Equality of Variance Matrices

Box's M	278.858
F	3.362
df1	75
df2	9170.116
Sig.	.000

Note. Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

Assumption 10: Homogeneity of variances. There should be homogeneity of variances. That is, there should be equal variances in each cell of the design for each dependent variable. I used the Levene's test of equality of covariance matrices to test Assumption 10. The result of the Levene's test shows that the assumption was met for three of the dependent variables (measurement, statistics, and RGF), and violated for two of the dependent variables. (algebra and geometry). According to Bathke et al. (2018), the assumption of homogeneity of variance is not reasonable for realistic data application. Hence, having met the assumption for three of the dependent variables, I proceeded with the analysis. The results of the test are indicated in Table 15.

Table 15

The Levene's Test of Equality of Variance

	F	df1	df2	Sig.
Algebra	1.723	7	361	.102
Geometry	1.558	7	361	.147
Measurement	7.786	7	361	.000
Statistics	4.234	7	361	.000
RFG	2.275	7	361	.028

Note. Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept + Comprehension + Knowledge + Reasoning + CoP + Comprehension * Knowledge + Comprehension * Reasoning + Comprehension * CoP + Knowledge * Reasoning + Knowledge * CoP + Reasoning * CoP + Comprehension * Knowledge * Reasoning + Comprehension * Knowledge * CoP + Comprehension * Reasoning * CoP + Knowledge * Reasoning * CoP + Comprehension * Knowledge * Reasoning * CoP.

The 2-way MANOVA has 10 assumptions that must be considered before conducting statistical analyses. The first three assumptions relate to the design of the study (Ates et al., 2019; Ito, 1980), and were met prior to conducting the study. The three assumptions were: there must be two or more dependent variables at the continuous level, there must be two or more independent variables with at least two sub-levels, there must be independence of observation. The remaining seven assumptions were tested prior to conducting the analysis. Three of the seven assumptions were met in full, while three assumptions were partially met (met for some, but not all, variables). The assumptions met in full were Assumptions 4, 5, and 8, which were linearity, assessed by scatterplots; multicollinearity, assessed by Pearson correlation ($|r| < 0.9$); and adequate sample size, assessed by descriptive statistics, respectively. For Assumption 6 univariate outlier, assessed by boxplots was met for some of the variables, however, multivariate outliers, indicated by Mahalanobis distance ($p > .001$), was met for all variables. Assumption 9, homogeneity of variance-covariance matrices, assessed by Box's M test ($p < .001$); and Assumption 10, homogeneity of variances; were met for three of the five dependent variables. Assumption 7, multivariate normality, was not met. Overall, only one of the 10 assumptions for the two-way MANOVA was not met.

Hypothesis Testing

MANOVA. A two-way MANOVA was conducted to test the null hypothesis that there are no differences in the CSEC mathematics scores between high-scoring Vincentian students and low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning, across the content domains of algebra, geometry, measurement, statistics, and RFG. The cognitive domains were the focal variables and CoP was the moderator variable. Since there was violation of the homogeneity of variances and covariances as assessed by Box M's test ($p < .001$), Pillai's Trace was interpreted. Pillai's Trace is a superior, robust omnibus MANOVA test (Olson, 1976). There was a statistically significant interaction effect between knowledge and reasoning on the combined dependent variables, $F(5, 357) = 3.50$, $p = .004$, Pillai's Trace = .047, partial $\eta^2 = .047$. There was no interaction effect between the other independent variables on the combined dependent variables. There was a significant main effect for each of the four independent variables. Knowledge $F(5, 357) = 12.925$, $p < .001$, Pillai's Trace = .153, partial $\eta^2 = .153$. Comprehension $F(5, 357) = 10.025$, $p < .001$, Pillai's Trace = .123, partial $\eta^2 = .123$. Reasoning $F(5, 357) = 5.633$, $p < .001$, Pillai's Trace = .073, partial $\eta^2 = .073$. CoP $F(5, 357) = 2.414$, $p < .001$, Pillai's Trace = .033, partial η^2 =

.033. The information is displayed in Table 16 and represent significant *p* values only. For the full MANOVA results, see Appendix L.

Table 16

MANOVA Multivariate Results

Effect		Value	*F*	Hypothesis df	Error df	Sig.	Partial eta squared
Intercept	Pillai's Trace	.721	184.672[b]	5.000	357.000	.000	.721
Knowledge	Pillai's Trace	.153	12.925[b]	5.000	357.000	.000	.153
Comprehension	Pillai's Trace	.123	10.025[b]	5.000	357.000	.000	.123
Reasoning	Pillai's Trace	.073	5.633[b]	5.000	357.000	.000	.073
CoP	Pillai's Trace	.033	2.414[b]	5.000	357.000	.036	.033
Knowledge * Comprehension	Pillai's Trace	.000	.[b]	.000	.000	.	.
Knowledge * Reasoning	Pillai's Trace	.047	3.500[b]	5.000	357.000	.004	.047
Comprehension * Reasoning	Pillai's Trace	.000	.[b]	.000	.000	.	.
Knowledge * CoP	Pillai's Trace	.000	.[b]	.000	.000	.	
Comprehension * CoP	Pillai's Trace	.000	.[b]	.000	.000	.	.
Reasoning* CoP	Pillai's Trace	.000	.[b]	.000	.000	.	.
Knowledge * Comprehension *Reasoning	Pillai's Trace	.000	.[b]	.000	.000	.	.
Knowledge * Comprehension* CoP	Pillai's Trace	.000	.[b]	.000	.000	.	.
Knowledge *Reasoning * CoP	Pillai's Trace	.000	.[b]	.000	.000	.	.
Comprehension * Reasoning * CoP	Pillai's Trace	.000	.[b]	.000	.000	.	.
Knowledge * Comprehension* Reasoning * CoP	Pillai's Trace	.000	.[b]	.000	.000	.	.

Note. a. Intercept + Knowledge + Comprehension + Reasoning + CoP + Knowledge * Comprehension + Knowledge * Reasoning + Comprehension * Reasoning + Knowledge * CoP + Comprehension * CoP + Reasoning* CoP + Knowledge * Comprehension *Reasoning + Knowledge * Comprehension* CoP + Knowledge *Reasoning * CoP + Comprehension * Reasoning * CoP + Knowledge * Comprehension* Reasoning * CoP
b. exact statistics.

ANOVA. As a follow up to the interaction effect between knowledge and reasoning, separate two-way ANOVAs, between subject analyses were conducted for each content domain to determine whether there was any statistically significant effect for each dependent variable separately (Pituch & Stevens, 2016). There was a statistically significant interaction effect between level of knowledge and level of reasoning for measurement scores, $F(1, 369) = 16.634$, $p < .001$, partial $\eta^2 = .044$. There was no significant interaction effect between knowledge and reasoning for the other four content domains: algebra, $F(1, 369) = 1.151$, $p = .284$, partial $\eta^2 = .003$, geometry, $F(1, 369) = .122$, $p = .727$, partial $\eta^2 = .000$, statistics, $F(1, 369) = .523$, $p = .470$, partial $\eta^2 = .001$, or RFG, $F(1, 369) = 1.506$, $p = .221$, partial $\eta^2 = .004$. The results of the ANOVA tests are presented in Table 17.

Table 17

Test of Between-Subject Effect

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	*F*	Sig.	Partial Eta Squared
Knowledge * Reasoning	Algebra	7.635	1	7.635	1.151	.284	.003
	Geometry	.782	1	.782	.122	.727	.000
	Measurement	135.100	1	135.100	16.634	.000	.044
	Statistics	3.768	1	3.768	.523	.470	.001
	RFG	9.122	1	9.122	1.506	.221	.004

I computed simple main effect of knowledge and reasoning separately using univariate test to determine the difference between high and low levels of knowledge and high and low levels of reasoning on the content domain, measurement. There was a statistically significant difference between high and low levels of knowledge on measurement scores, $F(1, 366) = 78.50$, $p < .001$, partial $\eta^2 = .177$. There was a statistically significant difference between high and low levels of reasoning on measurement scores, $F(1, 366) = 54.881$, $p < .001$, partial $\eta^2 = .130$. Data are mean $\pm$ standard deviation, unless otherwise stated. The means for measurement scores were 13.87 $\pm$ 3.76 for high levels of knowledge and 10.20 $\pm$ 2.75 for low levels of knowledge, and 9.38 $\pm$ 2.79 for high levels of reasoning and 4.77 $\pm$ 3.36 for low levels of reasoning. There was a statistically significant difference between high levels of knowledge and low levels of knowledge, 3.67 (95% CI, 3.855 to 6.054), $p < .0005$; and between high levels of reasoning and low levels of reasoning, 4.60 (95% CI, 3.04 to 5.24), $p < .0005$.
The results are illustrated in Tables 18 and 19.

Table 18

Univariate Tests: Levels of Knowledge and Measurement

	Sum of squares	Df	Mean square	*F*	Sig.	Partial eta squared
Contrast	889.819	1	889.819	78.499	.000	.177
Error	4148.748	366	11.335			

Note. *F* tests the effect of Levels of Knowledge. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

Table 19

Univariate Test: Levels of Reasoning and Measurement

	Sum of squares	Df	Mean square	*F*	Sig.	Partial eta squared
Contrast	622.102	1	622.102	54.881	.000	.130
Error	4148.748	366	11.335			

Note. *F* tests the effect of Levels of Reasoning. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

I conducted separate 2-way ANOVAs for each content domain to examine the main effects for knowledge, comprehension, reasoning, and CoP. There was a statistically significant difference between high and low levels of knowledge for four of the five content domain scores. For algebra scores, $F(1, 369) = 11.533$, $p = .001$, partial $\eta^2 = .031$; geometry scores, $F(1, 369) = 18.710$, $p < .001$, partial $\eta^2 = .049$; measurement scores, $F(1, 369) = 41.182$, $p < .001$, partial $\eta^2 = .102$; RFG scores $F(1, 369) = 15.36$, $p < .001$, partial $\eta^2 = .056$. The difference between high and low levels of knowledge was not statistically significant for statistics scores, $F(1, 369) = 1.719$, $p = .191$, partial $\eta^2 = .005$. The information is displayed in Table 20.

Table 20

Test of Significance: Levels of Knowledge and Content Domains

Dependent variable	*F*	Sig	Partial eta squared
Algebra	11.533	$p = .001$	.031
Geometry	18.710	$p < .001$	.049
Measurement	41.182	$p <. 001$	.102
Statistics	1.719	$p = .191$	.005
RFG	21.405	$p < .000$	.056

There was a statistically significant difference between high and low levels of comprehension for four of the five content domain scores. For algebra scores, $F(1, 369) = 15.096$, $p < .001$, partial $\eta^2 = .040$; geometry scores, $F(1, 369) = 22.189$, $p < .001$, partial $\eta^2 = .058$; measurement scores, $F(1, 369) = 24.791$, $p < .001$, partial $\eta^2 = .064$; RFG scores $F(1, 369) = 13.056$, $p < .001$, partial $\eta^2 = .035$. The difference between high and low levels of comprehension was not statistically significant for statistics scores, $F(1, 369) = 9.875$, $p = .243$, partial $\eta^2 = .004$. The information is displayed in Table 21.

Table 21

Test of Significance: Levels of Comprehension and Content Domains

Dependent variable	*F*	Sig	Partial eta squared
Algebra	15.096	$p < .001$	.040
Geometry	22.189	$p < .001$	.058
Measurement	24.791	$p < .001$	.064
Statistics	1.370	$p = .243$	.004
RFG	13.056	$p < .001$	.035

There was a statistically significant difference between high and low levels of reasoning for four of the five content domain scores. For algebra scores, $F(1, 369) = 9.649$, p = .002, partial $\eta^2 = .026$; geometry scores, $F(1, 369) = 10.171$, $p = .002$, partial $\eta^2 = .027$; measurement scores, $F(1, 369) = 14.479$, $p < .001$, partial $\eta^2 = .039$; RFG scores $F(1, 369) = 7.722$, $p = .006$, partial $\eta^2 = .021$. The difference between high and low levels of reasoning was not statistically significant for statistics scores, $F(1, 369) = .341$, p = .576, partial $\eta^2 = .001$. The information is displayed in Table 22.

Table 22

Test of Significance: Levels of Reasoning and Content Domains

Dependent variable	*F*	Sig	Partial eta squared
Algebra	9.649	p = .002	.026
Geometry	10.171	p = .002	.027
Measurement	14.479	p < .001	.039
Statistics	.314	p = .576	.001
RFG	7.722	p = .006	.021

There was a statistically significant difference between high and low CoP for two of the five content domain scores. For algebra scores, $F(1, 369) = 9.815$, $p = .002$, partial $\eta^2 = .026$; RFG scores, $F(1, 369) = 5.359$, $p = .021$, partial $\eta^2 = .015$. The difference between CoP groups was not statistically significant for geometry scores, $F(1, 369) = 3.475$, $p <= .063$, partial $\eta^2 = .010$; measurement scores $F(1, 369) = .588$, $p = .444$, partial $\eta^2 = .002$; statistics scores, $F(1, 369) = 1.252$, $p = .021$, partial $\eta^2 = .015$. The information is displayed in Table 23.

Table 23

Test of Significance CoP and Content Domains

Dependent variable	*F*	Sig	Partial eta squared
Algebra	9.815	p = .002	.026
Geometry	3.475	p =.063	.010
Measurement	.588	p =.444	.002
Statistics	1.252	p= .264	.003
RFG	5.359	p =.021	.015

Answering the Research Question

The primary research question that guided this study was "How do the CSEC mathematics scores of high-scoring Vincentian students versus low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning differ across the content domains of algebra, geometry, measurement, statistics, and RFG?". The related null hypothesis was "There are no differences in the CSEC mathematics scores between high-scoring Vincentian students and low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning, across the content domains of algebra, geometry, measurement, statistics and RFG".

I used 2-way MANOVA and follow up 2-way ANOVA univariate analysis to test the null hypothesis. The multivariate results of Pillai's Trace indicated a statistically significant interaction effect between knowledge and reasoning on the combined dependent variables, $p = .004$. The null hypothesis that the combination of algebra, geometry, measurement, statistics, and RFG is the same for all combinations of high and low levels of knowledge, and high and low levels of reasoning was therefore rejected. I then conducted follow up ANOVA univariate to further test the hypothesis to determine which content domain(s) contributed to the statistically significant interaction effect shown between levels of knowledge and levels of reasoning. The results of the univariate analysis indicated that the content domain of measurement, $p < .001$, was responsible for the observed interaction effect. Separate univariate tests for levels of knowledge and levels of reasoning for measurement indicated that levels of knowledge explained 17.7 % of the variance in the interaction, whereas levels of reasoning explained 13%.

Statistically significant simple main effects were observed for knowledge, $p < .001$, comprehension, $p < .001$, reasoning, $p < .001$, and CoP, p =. 036. Follow up 2-way ANOVA was conducted to determine which content domains were responsible for the observed main effects among each independent variable. The difference between high and low levels of knowledge was found to be significant for algebra, $p = .001$, geometry, $p < .001$; measurement, $p < .001$, and RFG, $p < .001$. The difference between high and low levels of knowledge was not statistically significant for statistics, $p = .191$. The difference between high and low levels of comprehension was found to be significant for algebra, $p < .001$; geometry, $p < .001$; measurement, $p < .001$, and RFG, $p < .001$. The difference between high and low levels of knowledge was not statistically significant for statistics, $p = .243$. The difference between high and low levels of reasoning was found to be significant for algebra, $p = .002$; geometry, $p = .002$; measurement, $p < .001$, and RFG, $p = .006$. The difference between high and low levels of knowledge was not statistically significant for statistics, $p = .576$. The difference between high and low CoP was found to be significant for algebra, $p = .002$, and RFG, p < .021. The difference between high and low CoP was not statistically significant for geometry, $p = .663$; measurement, $p = .444$; and statistics, $p = .264$.

Summary

The focus of Chapter 4 was the presentation of the results of the study. The chapter began with an overview of the study, including the purpose, research question, and hypotheses. Following this, I described the data collection procedures and presented the results, including the outcomes of the assumptions and the statistical tests used to answer the research questions. Highlights of the findings are summarized below.

I conducted 2-way MANOVA and follow up 2-way ANOVA tests to examine the relationship between the cognitive domains of knowledge, comprehension and reasoning, and CoP, and the content domains of algebra, geometry, measurement, statistics, and RFG. I used the composite content domain scores to determine CoP. Prior to conducting the analyses, I tested the assumptions of the 2-way MANOVA. Overall, one of the 10 assumptions for the 2-way MANOVA was violated.

The results of the 2-way MANOVA indicated a statistically significant interaction effect between levels of knowledge and levels of reasoning on the combined dependent variables $F(1, 369) = 16.634$, $p < .001$, partial $\eta^2 = .044$. There was no interaction effect between the other independent variables on the combined dependent variables. Follow up univariate 2-way ANOVA tests indicated a statistically significant interaction effect between levels of knowledge and levels of reasoning for measurement scores, $F(1, 369) = 16.634$, $p < .001$, partial $\eta^2 = .044$. There was a significant main effect for each of the four independent variables. Knowledge $F(5, 357) = 12.925$, $p < .001$, Pillai's Trace = .153, partial $\eta^2 = .153$. Comprehension $F(5, 357) = 10.025$, $p < .001$, Pillai's Trace = .123, partial $\eta^2 = .123$. Reasoning $F(5, 357) = 5.633$, $p < .001$, Pillai's Trace = .073, partial $\eta^2 = .073$. CoP $F(5, 357) = 2.414$, $p < .001$, Pillai's Trace = .033, partial $\eta^2 = .033$.

I computed simple comparisons for differences in mean measurement scores between high and low levels of knowledge. Data are mean $\pm$ standard deviation, unless otherwise stated. The means for measurement scores were 13.87 ± 3.76 for high levels of knowledge and 10.20 ± 2.75 for low levels of knowledge, and 9.38 ± 2.79 for high levels of reasoning, and 4.77 ± 3.36 for low levels of reasoning. There was a statistically significant difference between high levels of knowledge and low levels of knowledge, 3.67 (95% CI, 3.855 to 6.054), $p < .0005$; and between high levels of reasoning and low levels of reasoning, 4.60 (95% CI, 3.04 to 5.24),), $p < .0005$. Overall, measurement scores were responsible for the significant interaction effect between knowledge and reasoning. Significant main effects were noted for algebra, geometry, measurement, and RFG in each cognitive domain, and in algebra and RFG for CoP. In Chapter 5, I interpret the findings of the study, discuss the limitations, outline recommendations, and discuss the implications of the study for social change, and educational practice.

Chapter 5: Discussion, Conclusions, and Recommendations

Introduction

The purpose of this quantitative, nonexperimental study was to determine the extent to which the CSEC mathematics scores of high-scoring Vincentian students versus low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning differ across five content domains of algebra, geometry, measurement, statistics, and RFG. The secondary data used in the study comprised the scores of Vincentian students in the 2017 May/June CSEC mathematics examination. The study was conceptualized and designed in response to the poor performance of Vincentian students in the CSEC mathematics examination, and the particularly poor performance on the cognitive domain of reasoning. I conducted a 2-way MANOVA to examine the relationship between the independent variables of knowledge, comprehension, reasoning, and CoP and the dependent variables of algebra, geometry, measurement, statistics, and RFG.

In this chapter, I interpret and discuss the findings of the study. The discussion includes how the findings relate to the literature review synthesized in Chapter 2 and in the context of the theoretical framework. The limitations of the study follow the interpretation of the findings and include the generalizability of the findings. Recommendations for further research are then presented in the context of the strengths and limitations of the study and the current literature in the field. I outline implications of the study for positive social change and educational practice. The chapter culminates with a conclusion that summarizes the highlights of the study including the purpose, findings, and implications.

Interpretation of the Findings

In this section, I discuss the interpretations of the findings in relation to current literature in the field and the theoretical framework that guided the study.

Mathematics competency is critical to daily existence and efficient functioning in modern societies (Bosman & Schulze, 2018; Hassan et al., 2019; Primi et al., 2020; Waxman, 2020). Mathematics is a multidimensional construct that encompasses different cognitive skills and abilities, as well as cognitive and noncognitive factors that have been found to play a significant role in mathematics achievement (Cirino, Tolar, Fuchs, & Huston-Warren, 2016; Cormier et al., 2017; Gilmore et al., 2018; Männamaa et al., 2012; O'Connell, 2018; Passolunghi, Cargnelutti, & Pellizzoni, 2019; Semeraro et al., 2020; Skagerlund & Träff, 2016). Although the relationship between cognitive abilities and mathematics achievement has been well established, some studies have produced mixed or opposite results (Areepattamannil & Caleon, 2013; Caemmerer et al., 2018). Cormier et al. (2017) suggested that for studies that reported weak relationships between cognitive abilities and mathematics achievement, a number of factors may be responsible for such differences. These factors may include the specific area of mathematics that was investigated, the different components of cognitive abilities examined, or the lack of a common nomenclature used to identify the cognitive abilities.

Interpretation of Findings in Relation to Current Literature

In the present study, there was a statistically significant interaction effect between two cognitive domains: levels of knowledge and levels of reasoning on the combined dependent variables of algebra, geometry, measurement, statistics, and RFG, $F(5, 357) = 3.50$, $p = .004$, Pillai's Trace = .047, partial $\eta^2 = .047$. The interaction effect was significant for measurement scores only, $F(1, 369) = 16.634$, $p < .001$, partial $\eta^2 = .044$. These findings indicate that the measurement scores of students in the high knowledge group were statistically significantly higher than those of their counterparts in the low knowledge group. Similarly, the measurement scores of students in the high-reasoning group were statistically significantly higher than those of their counterparts in the low-reasoning group. There was no interaction effect between the other independent variables, namely, knowledge and comprehension, comprehension and reasoning, knowledge and CoP, comprehension and CoP, and reasoning and CoP, on the combined dependent variables. The results of the MANOVA led to the rejection of the null hypothesis that there are no differences in the CSEC mathematics scores between high-scoring Vincentian students and low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning across the content domains of algebra, geometry, measurement, statistics, and RFG.

There was a significant main effect for each of the four independent variables. Knowledge $F(5, 357) = 12.925$, $p < .001$, Pillai's Trace = .153, partial $\eta^2 = .153$. Reasoning $F(5, 357) = 5.633$, $p < .001$, Pillai's Trace = .073, partial $\eta^2 = .073$. CoP $F(5, 357) = 2.414$, $p < .001$, Pillai's Trace = .033, partial $\eta^2 = .033$. These findings are consistent with the findings of earlier studies (Caemmerer et al., 2018; Cormier et al., 2017; Cowan et al., 2018; Floyd et al., 2003; Green et al., 2017; Khan, 2017; Primi et al., 2010; Taub et al., 2008). Green et al. (2017), Gelbart (2007) and Primi et al. (2010) found fluid reasoning to be a significant predictor of student mathematics reasoning. Similar to the present study, the mathematics domain in the Primi et al. (2010) study included geometry, numbers, functions, and statistics. In the present study, students in the high-knowledge group, high-comprehension group, and high-reasoning group performed significantly better than their counterparts in the low-knowledge group, low-comprehension group, and low-reasoning group on all content domains. These findings support earlier findings by Primi et al. (2010) in which they reported that students with higher intelligence showed faster increases in mathematics scores than their counterparts with lower fluid intelligence. The findings of Primi et al. (2010) as well as those of the present study were further corroborated by the findings of a meta-analysis by Peng et al. (2019) in which they reported a strong reciprocal relationship between fluid intelligence and mathematics. In a similar study which investigated working memory capacity and student performance in arithmetic, percentages, proportion, decimals, algebra, and geometry, Musso et al. (2019) found a direct relationship between working memory capacity and mathematics performance. The findings of the current study support earlier research by Lee and Bull (2016) who found that students with higher working memory or updating capacity performed better than their counterparts with lower working memory of updating capacity. Other researchers who investigated the relationship between specific cognitive abilities and achievement in specific mathematics content domains reported a moderate to strong relationship between comprehension knowledge and mathematics achievement particularly in the areas of mathematics problem solving and mathematics calculation skills (Cormier et al., 2017; Floyd et al., 2003; & Taub et al., 2008). Passolunghi et al. (2019) also found student working memory and processing speed to be strongly associated with high performance on arithmetic problem solving. Overall, the findings of the present study support earlier research in the field of cognitive abilities and mathematics achievement. The findings also extend current literature by adding the influence of cognitive abilities on select mathematics content domains.

Interpretation of Findings in Context of the Theoretical Framework

The theoretical perspective of the study was based on Bloom’s taxonomy of educational objectives. Bloom’s taxonomy of educational objectives (Bloom et al., 1956) is a pedagogical tool designed to guide educators in developing meaningful assessment of learning outcomes (Ramirez, 2017). The taxonomy has filled a void by providing a basis by which educators can systematically evaluate students’ learning (Bertucio, 2017; Hadzhikoleva et al., 2019). Bloom’s taxonomy is debatably one of the most prominent educational monographs produced in the last 5 decades (Cullinane & Liston, 2016). Bloom’s taxonomy was chosen as the theoretical base for this study because of its simplistic nature which allows for certainty and efficiency with which higher-order questions and lower-order questions could be distinguished (Cullinane & Liston, 2016).

Bloom’s taxonomy provides a framework for assessment including a model for identifying the cognitive processes examiners use when solving problems (Bloom et al., 1956). Given established purpose, the CXC used Bloom’s taxonomy as the framework for the development of the CSEC mathematics examination. The CXC used the cognitive processes defined by the taxonomy to develop a cognitive by content matrix that formed the blueprint for the test. Achievement in the CSEC mathematics examination is consistent with Männamaa et al. (2012) view of mathematics achievement, which they perceive as comprising two dimensions: the content of the task, which includes the topics, and the cognitive abilities needed for solving these tasks such as knowing, computing, knowing and using algorithms, solving word problems, and applying these skills in novel situations (Männamaa et al., 2012). A thorough understanding of mathematics achievement requires an identification of important relationships between cognitive skills and specific components of mathematics (Gilmore et al., 2018). I sought to contribute to the literature on the influence of cognitive abilities on mathematics achievement by analyzing students' performance in the CSEC mathematics examination by cognitive domain based on three levels of Bloom's taxonomy, knowledge, comprehension, and reasoning, and five mathematics content domain: algebra, geometry, measurement, statistics, and RFG.

The results of the present study showed that students in the high-knowledge group and the high-reasoning group had statistically significant higher scores on the measurement domain than their counterparts in the low-knowledge and low-reasoning group. The study's findings further corroborate the findings of other researchers who found that specific cognitive domains predict achievement in specific mathematics domain (Khan, 2017; Männamaa et al., 2012; & Zhang et al., 2017). Khan (2017) found that although students in the national Grade 4 mathematics test in Trinidad and Tobago generally performed poorly on the measurement domain, students in the lower-performing group have statistically significantly lower scores than their counterparts in the higher-performing group.

The findings of the present study also indicated that students in the high-knowledge group, the high-comprehension group, and high-reasoning group had significantly higher scores on algebra, geometry, measurement, and RFG, than their counterparts in the low-knowledge group, the low-comprehension group, and low-reasoning group. The scores in statistics were not significantly different for students in the high group and the low group for knowledge, comprehension, and reasoning. Additionally, students in the high-CoP group had significantly higher scores in algebra and RFG than their counterparts in the low-CoP group. These findings are consistent with those of earlier studies that found cognitive domain to be a significant predictor of student mathematics achievement (Primi et al., 2010; Vista, 2016; Wong & Ho, 2017). The findings also support the view that tasks requiring higher cognitive skills improve critical thinking skills and result in more permanent learning (Tarman & Kuran, 2015).The findings of the present study also support those of Khan (2017) who reported that overall, questions which required higher-order thinking skills posed the greatest challenges for all students. According to O'Connell (2018), cognitive ability is a key driver of academic achievement for most students. Also, given the role of fluid intelligence, reasoning, and problem solving in predicting mathematics reasoning (Cowan et al., 2018; Green, et at., 2017; Passolunghi et al., 2019; Primi et al., 2010; Semeraro et al., 2020), it is not surprising that in this study, students in each of the high-cognitive groups of knowledge, comprehension, and reasoning had significantly higher mathematics scores than their counterparts in each of the low-cognitive groups.

Specific cognitive abilities have been found to predict achievement in specific mathematics domains (Gilmore et al., 2018; Khan, 2017; Männamaa et al., 2012; Zhang et al., 2017). In the present study, the scores of students in the high group for knowledge, comprehension, and reasoning were statistically significantly higher than those of their counterparts in the low groups, for four content domains, algebra, geometry, measurement, and RFG, but not for statistics. Students scored lowest in the content domains of measurement and statistics. The findings of the study partially support those of Khan (2017) who reported that students achieved their highest scores in statistics and their lowest scores in measurement. It may be important to note that Khan's study included Grade 6 students, whereas the present study included Grade 11 students. A possible explanation for the lack of difference between the high group and the low group for each of the cognitive domains in the present study may be related to the type of tasks students were required to perform for the statistics items, and whether the cognitive levels for the tasks were accurately and consistently differentiated by the test developers.

For CoP, the scores of the students in the high groups were significantly higher than those of the students in the low group for algebra and RFG. However, the scores of the students in the high group and the low group were not significantly different for geometry, measurement, and statistics. These results suggest that students demonstrated different levels of competence in the various mathematics domains. Based on this outcome, it is very likely that students with the same composite scores did not have the same pattern of scores for geometry, measurement, and statistics, resulting in some students in the low CoP group having higher scores than some students in the high CoP group on geometry, measurement, and RFG.

The overall findings of the study support the view that cognitive abilities have a significant influence on student mathematics achievement as demonstrated by students in the high-scoring groups in each cognitive domains: knowledge, comprehension, and reasoning scoring significantly higher than their counterparts in the corresponding low-scoring groups in four of the five content domains investigated. The outcomes of the study are consistent with the purpose of Bloom's taxonomy in providing educators with a framework for understanding learning outcomes and delineating tasks involving higher and lower-order skills (Irvine, 2017), as well as providing clarity in designing and sequencing educational objectives (Ramirez, 2017).

The conceptual model in Figure 5 summarizes the main outcomes of the study, supporting the influence of cognitive abilities, as defined by Bloom's taxonomy, on mathematics achievement. The conceptual model gives a diagrammatical illustration of the outcomes of the study. It highlights the interaction effect which occurred between levels of reasoning and levels of knowledge, and the main effects of each cognitive domain: knowledge, comprehension, and reasoning, on the content domains. In addition to providing a summary of the primary findings, the model consolidates the relationship between cognitive abilities and achievement of Vincentian students in the 2017 May/June CSEC mathematics examination. This model may be used as a foundation for further exploration of the role of cognitive abilities in mathematics achievement in other contexts.

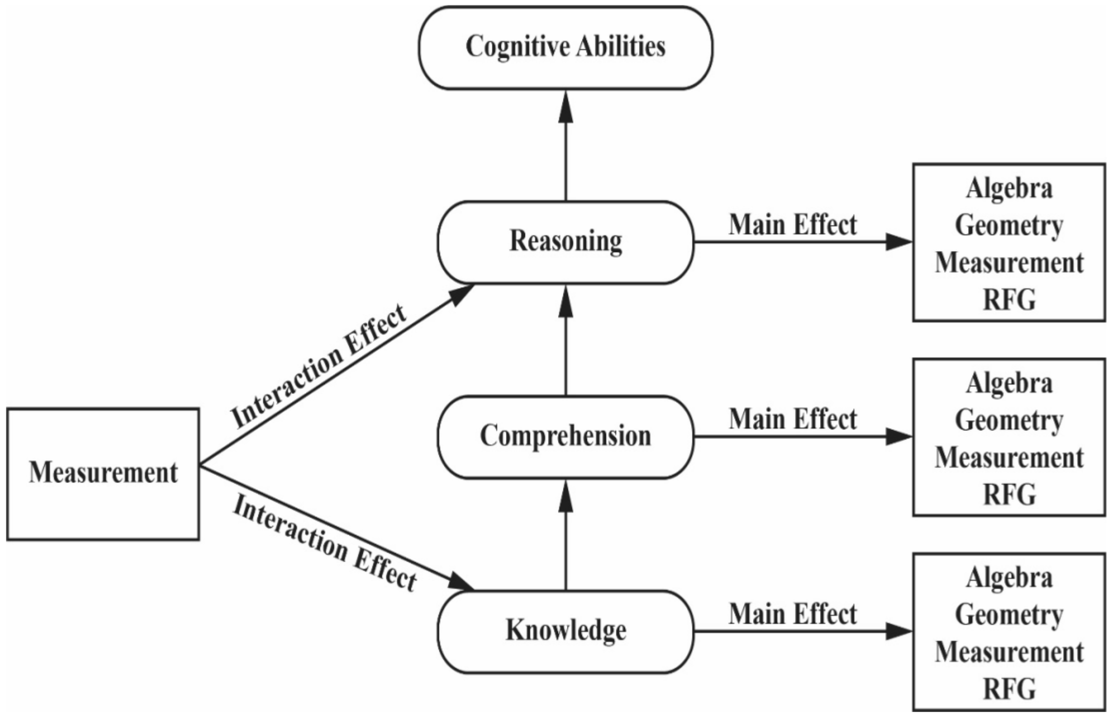


Figure 5. Conceptual model: Mathematics cognitive abilities (Bloom's taxonomy) and Content Domains in CSEC Mathematics.

Limitations of the Study

The study was nonexperimental with a cross-sectional designed that used archival data. The purpose of the study was to determine whether the CSEC mathematics scores of high-scoring Vincentian students versus low-scoring Vincentian students in the cognitive domains of knowledge, comprehension, and reasoning differ across the content domains of algebra, geometry, measurement, statistics, and RFG. Limitations of the study include the cross-sectional design and the use of archival data. The cross-sectional design makes it difficult to make causal inferences (Bono & McNamara, 2011; Levin, 2006).

The data used in the study were collected at one point in time and comprised students' mathematics examination scores for 1 year. Data on student performance for a single examination does not allow for the establishing a trend in performance. Basing analyses on data collected at a single-point-in-time limits the generalizability of the results to other examination sittings. Archival data are not always organized in a form that will allow for maximum data usage. For instance, the data received included students' cognitive domain scores and content domain scores. The gender and age of the students were not included, which could have been used for further exploration. Other limitations included the use of five content domains with total scores ranging from 16 to 20 and scores on the cognitive domains ranging from 28 to 34. A maximum of 20 marks or fewer for a content domain may not have been sufficient to give adequate content coverage for the domains to give an accurate determination of students' proficiency level in the particular content domain, and establish content validity of the examination. Another limitation identified was that the cognitive domains may not have be consistently operationalized in all content domains throughout the examination. Accurate and consistent operationalization of cognitive domains is required to establish sound psychometric properties of the examination and produce reliable scores from which to draw valid inferences regarding student performance.

Recommendations

The results of the present study contribute to the growing literature on cognitive abilities and mathematics abilities by exploring a unique combination of content domains using the cognitive abilities as defined by Bloom taxonomy of learning objectives

(Bloom et al., 1956). The study is the first of its kind to be conducted in the Caribbean and Saint Vincent and the Grenadines, in particular. The findings of the study support previous research studies in demonstrating a strong relationship between cognitive abilities and mathematics achievement. Based on the findings of the study, I offer recommendations for further research and practice.

Recommendations for Further Research

Given the importance of mathematics to overall academic achievement, and in light of the general poor performance of students in all content domains, future research may investigate the type of strategies teachers use when teaching mathematics. In many classrooms in Saint Vincent and the Grenadines, "chalk and talk" is still the predominant mode of teaching. This method of teaching encourages rote learning among students. In most cases, the teacher is the focal point of attention and dictates the solutions of mathematics problems. Hence, students may not be afforded sufficient opportunities to discover solutions for themselves. This type of rote learning does not foster critical thinking which is necessary for developing higher order cognitive skills, such as reasoning, which has been found to be related to mathematics achievement (Cormier et al., 2017; Cowan et al., 2018; Semeraro et al., 2020).

Researchers may also explore the type of strategies students use when solving mathematics problems. Some students seem to rely on a surface approach, in which they simply try to recall facts and procedure, as opposed to a deep approach in which they are engaged with the material and able to apply their knowledge to new situations. The surface approach to learning is consistent with the knowledge level of Bloom's taxonomy, which is used to guide the construction of the CSEC mathematics examination. The outcomes of the study have shown that students in the high-reasoning group had higher mathematics scores, hence future research may investigate the effect of a surface approach versus a deep approach in solving mathematics problems.

Researchers may also explore curriculum alignment of the CSEC mathematics examination to determine whether there is congruence among the students' expectations, instruction, and assessment. Nonalignment of curriculum has been identified as one of the reasons for students' poor performance (Bhaw & Kriek, 2020; Seitz, 2017; Squires, 2012). Exploration of curriculum alignment may include examining balance between representation and cognitive complexity. Balance of representation focusses on topic coverage between the curriculum and assessment, whereas cognitive complexity focusses on cognitive demand between curriculum and assessment (Bhaw & Kriek, 2020).

Researchers may also explore the influence of cognitive abilities on mathematics achievement using other content domains as the measure of achievement. Additionally, research may include other examination years to explore whether there is a trend in performance. Future research may also consider replicating this study in other Caribbean territories to determine whether the findings will hold true. Researchers may also investigate the influence of cognitive abilities on mathematics achievement using gender, age, school type, school location, or socioeconomic status as mediating variables.

Recommendations for Practice

Insights into the influence of cognitive abilities on student mathematics achievement may help education administrators identify students who are at risk of developing learning difficulties in mathematics so that they could plan intervention strategies for remediation to improve student mathematics achievement.

Teachers can help students to improve their cognitive abilities by ensuring that classroom instructions and assessments are aligned to the syllabus and examination specifications based on the CXC's requirements for the CSEC mathematics examination. Teachers are therefore encouraged to use Bloom's taxonomy of educational objectives to model their classroom pedagogy. They can do so by using the specimen paper, which is a blueprint for the examination, as well as past examination papers to guide and develop their classroom assessments in which they could challenge students to engage in higher order thinking. Teachers may support students in developing higher cognitive abilities by challenging them to explore multiple solution to mathematic problems, and connecting procedures and concepts (Kieran, 2013; Star et al., 2015).

Creating a classroom culture that fosters the development of critical thinking skills through the use of skillful questioning that allows students to hone these skills is likely to result in higher academic achievement in high-stakes examination for students (Whittle, Benson, Ullah, & Telford, 2018). In many classrooms in Saint Vincent and the Grenadines, teachers still utilize behaviorist approaches in the teaching of mathematics, where the teacher is considered the custodian of knowledge and is responsible for transmitting that knowledge to the students (Ampadu & Danso, 2018). In the behaviorist approach to teaching, students are considered *tabula rasa*, empty vessels to be filled (Tirza, 2020). The behaviorist approach to teaching has been criticized for producing students who are unable to engage in critical thinking (Boaler & Staples, 2008), and to transfer knowledge acquired in mathematics classrooms to solving real-life problems (Ampadu & Danso, 2018). According to Lambert et al. (as cited in Ampadu & Danso, 2018), teachers need to adopt a more effective approach to pedagogy such as the constructivist approach that incorporates the learner's experiences, beliefs, and world views into the learning process. In using a constructivist approach, students will be actively involved in the learning process (Tirza, 2020), and would be challenged to construct knowledge from within (Tarman & Kuran, 2015).

With regard to the non-significant difference in the scores between high CoP and low CoP for geometry, measurement, and statistics, the CXC could seek to explore the content validity of these examination questions by commissioning mathematics content specialists and psychometricians to engage in question review and analysis, where necessary. The review and analysis should result in improvement in the psychometric properties of questions to be used in future examinations, including ensuring that the cognitive levels of questions are accurately assigned, and are commensurate with the requirements of the tasks.

Implications for Positive Social Change

The findings of the present study contribute to positive social change by providing teachers, administrators, and education policy makers in Saint Vincent and the Grenadines with insights into the influence of cognitive abilities on student achievement in the CSEC mathematics examination, including influence on specific content domains. With such insights, education administrators are better able to plan intervention strategies to help students to enhance their higher order cognitive skills, which is likely to improve students' mathematics competence. Strategies for enhancing cognitive skills may include adopting a constructivist approach to teaching that is more student oriented as well as pre-service and in-service teacher training to teach teachers to write higher-order questions that will challenge students to engage in critical thinking. Tasks that stimulate students' cognitive abilities motivate them to be fully engaged in the learning process, take more responsibility for their learning, and result in more permanent learning (Tarman & Kuran, 2015). Students who are mathematically proficient are likely to function more efficiently in society, lead more successful lives, and have better career options (Algarni, 2018; Dobie, 2019; Hassan et al., 2019; Primi et al., 2020; Waxman, 2020).

Conclusion

Mathematics achievement is a major component of student overall academic achievement and is critical to their effective functioning in a dynamic society that is becoming increasing quantified. However, students at all levels continue to demonstrate a lack of mathematics competence. The under achievement of students in mathematics is of great concern to education stakeholders globally, and has captured the attention of researchers who continue to seek reasons for such under achievement. Researchers have found a number of cognitive and noncognitive factors to be associated with mathematics achievement at various levels (Areepattamannil, & Caleon, 2013; Cowan et al., 2018; Semeraro et al., 2020; Xenidou-Dervou et al., 2018).

In this study, I sought to determine the effect of students' cognitive abilities, as defined by Bloom's taxonomy, on achievement in select mathematics content domains among a sample of Vincentian students on the 2017 May/June CSEC mathematics examination. The findings of the study indicated significant interaction between students in the high knowledge group and the low knowledge group, and between students in the high reasoning group and the low reasoning group on the measurement domain. There was also a significant main effect between students in the high-cognitive domain groups and the low-cognitive domain groups for algebra, geometry, measurement, and RFG. Additionally, the scores of students in the high-scoring group on the overall examination, were not significantly different from those in the low-scoring group on geometry, measurement, and statistics. Overall, the findings of the study support earlier studies that found cognitive abilities to play a significant role in mathematics achievement (Cormier et al., 2017; Cowan et al., 2018; O'Connell, 2018; Roth et al., 2015; Semeraro et al., 2020). Further research in the field of cognitive abilities and mathematics achievement are recommended. Such research may focus on cognitive abilities as defined by Bloom's taxonomy and may include different mathematics content domains, as well as students' data from other Caribbean territories. The research could also focus on performance trends and include student performance data across several years.

The findings of this study provide insights into the influence of Vincentian students' cognitive abilities on their mathematics achievement. Given the significance of students' reasoning abilities in their mathematics success, educators in Saint Vincent and the Grenadines may use the insights from this study to transform mathematics pedagogy with a view to improving the overall mathematics achievement of Vincentian students. Educators may adopt strategies such as student-centered approaches to instruction and open-ended classroom assessments that will challenge students to develop critical thinking and problem-solving skills that are critical to success in mathematics. Contemporary visions of improving mathematics achievement in the United States focus on teacher training that promotes student-centered learning and the solving of authentic problems (Ferrini-Mundy, 2000; Kieran, 2013; Litke & Corven, 2019). The findings of this research align with a student-centered approach and supports arguments for using a similar strategy to transform and support mathematics education efforts in Saint Vincent and the Grenadines. It is hoped that if implemented, these research-informed recommendations will help to produce mathematically competent students who will have greater access to higher paying career options and greater economic security, thereby positioning them to contribute in meaningful ways to their communities and society at large.

References

Abu Tayeh, K., Mohammad, S. A. R., & Mohammad, F. A. S. (2018). The reasons for the decline of the results of Jordanian students in "TIMSS 2015." *International Journal of Instruction, 11*(2), 325–338. doi:10.12973/iji.2018.11222a

Adams, N. E. (2015). Bloom's taxonomy of cognitive learning objectives. *Journal of Medical Library Association, 103*(3), 152–153. doi:10.3163/1536-5050.103.3.010

Adamson, B., Forestier, K., Morris, P., & Han, C. (2017). PISA, policymaking and political pantomime: Education policy referencing between England and Hong Kong. *Comparative Education, 53*(2), 192–208. doi:10.1080/03050068.2017.1294666

Ajello, A. M., Caponera, E., & Palmerio, L. (2018). Italian students' results in PISA mathematics test: Does reading competence matter? *Scholarly Journal*, 505–520. doi:10.1007/s10212-018-0385-x

Algarni, B. (2018). A meta-analysis on the effectiveness of flipped classroom in mathematics education. Paper presented at the Tenth International Conference on Education in New Learning Technologies. doi:10.21125/edulearn.2018.1852

Almerico, G. M., & Baker, R. K. (2004). Bloom's Taxonomy illustrative verbs: Developing a comprehensive list for educator use. *Florida Association of Teacher Educators Journal, 1*(4), 1–10. http://www.fate1.org/journals/2004/almerico1.pdf

Ampadu, E., & Danso, A. (2018). Constructivism in mathematics classrooms: Listening to Ghanaian teachers' and students' views. *Africa Education Review, 15*(3), 49–71. doi:10.1080/18146627.2017.1340808

Anderson, J. O., Lin, H. S., Treagust, D. F., Ross, S. P., & Yore, L. D. (2007). Using large-scale assessment datasets for research in science and mathematics education: Programme for International Student Assessment (PISA). *International Journal of Science and Mathematics Education, 5*, 591–614. doi:10.1007/s10763-007-9090-y

Anderson, L. W., Krathwohl, D. R., & Airasian, P. W. (Eds.). (2001). *A taxonomy of learning, teaching, and assessing*: *A revision of Bloom's Taxonomy of educational objectives*. New York, NY: Longman. Retrieved from https://nsee.memberclicks.net/assets/docs/KnowledgeCenter/EnsuringQuality/BooksReports/146.%20a%20taxonomy%20for%20learning.pdf

Areepattamannil, S., & Caleon, I. S. (2013). Relationships of cognitive and metacognitive learning strategies to mathematics achievement in four high-performing East Asian education systems. *The Journal of Genetic Psychology, 174*(6), 696–702. doi:10.1080/00221325.2013.799057

Atasoy, E. (2019). Elementary mathematics teachers' geometric thinking levels and their self-efficacy in geometry. *Acta Didactica Napocensia, 12*(2), 161–170. doi:10.24193/adn.12.2.12

Ates, C., Kaymanz, O. Z., Kale, H. E., & Tekindal, M. A. (2019). Comparison of test statistics of nonnormal and unbalanced samples for multivariate analysis of variance in terms of type-I error rates. *Computational and Mathematical Methods in Medicine*. doi:10.1155/2019/2173638

Balázsi, I., & Szepesi, I. (2018). Comparing results of TIMSS and the Hungarian national assessment of basic competencies. *Orbis Scholae, 12*(2), 65–76. doi:10.14712/23363177.2018.293

Bangdiwala, S. I. (2019). Basic epidemiology research design 1. Cross-sectional design. *International Journal of Injury Control and Safety Promotion, 1*, 124–126. doi:10.1080/17457300.2018.1556415

Barbieri, C. A., Miller-Cotto, D., & Booth, L. (2019). Lessening the load of misconceptions: Design-based principles for algebra learning, *Journal of the Learning Sciences*, *28*(3), 381–417. doi:10.1080/10508406.2019.1573428

Bassey, S. W., Joshua, M. T., & Asim, A. E. (2009, February). Gender differences and mathematics achievement of rural senior secondary students in Cross River State, Nigeria. *Proceeding of International Conference to Review Research in Science, Technology and Mathematics Education, (epiSTEME-2)*, Mumbai, India.

Bathke, A. C., Friedrich, S., Pauly, M., Konietschke, F., Staffen, W., Strobl, N., & Höller, Y. (2018) Testing mean differences among groups: Multivariate and repeated measures analysis with minimal assumptions. *Multivariate Behavioral Research, 53*(3), 348–359. doi:10.1080/00273171.2018.1446320

Battista, M. T. (2007). Development of geometric and spatial thinking. In F. L Lester, Jr. (Ed.). *Second handbook of research on mathematics teaching and learning* (pp. 843–908). Reston, VA: NCTM.

Berg, C., & Boote, S. (2017). Effects of empirically derived multiple-choice versus free-response instruments when assessing graphing abilities. *International Journal of Science & Mathematics Education, 15*(1), 19–38. doi:10.1007/s10763-015-9678-6

Bertucio, B. (2017). The cartesian heritage of Bloom's taxonomy. *Studies in Philosophy of Education, 36*, 477–497. doi:10.1007/sl1217-9575-2

Bhaw, N., & Kriek, J. (2020). The alignment of the grade 12 physics examination with CAPS curriculum: (November 2014–M 2018). *South African Journal of Education, 40*(1), 1–9. doi:10.15700/sage.v40n1a1676

Bjork, I. M., & Bowyer-Crane, C. (2013). Cognitive skills used to solve mathematical word problems and numerical operations: A study of 6– to 7–year-old children. *European Journal of Psychology of Education, 28*(4), 1345–1360. doi:10.1007/s10212-012-0169-7

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals: Handbook I: Cognitive domain.* New York, NY: David McKay.

Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside school. *Teachers College Record, 110*(3), 608–645.

Bono, J. E., & McNamara, G. (2011). Publishing in AJM–Part 2: Research design. *Academy of Management Journal, 54*(4), 657–660. doi:10.5465/amj.2011.64869103

Bosman, A., & Schulze, S. (2018). Learning style preferences and Mathematics achievement of secondary school learners. *South African Journal of Education, 38*(1), 1–8. doi:10.15700/saje.v38n1a1440

Braden, J. P. (2002). Best practices for school psychologists in educational accountability: High stakes testing and educational reform. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp.301–320). National Association of School Psychologists.

Breakspear, S. (2012). The policy impact of PISA: An exploration of the normative effects of international benchmarking in school system performance. *OECS Education Working Papers, 71*(1). doi.org/10.1787/5k9fdfqffr28-en

Brochu, P., Deussing, M. A., Houme, K., & Chuy, M. (2013). *Measuring up: Canadian results of the OECD PISA study, 2012 first results for Canadians aged 15.* Retrieved from http://cmec.ca/Publications/Lists/Publications/Attachments/318/PISA2012_CanadianReport_EN_Web.pdf

Browning, C., Edson, A. J., Kimani, P., & Asian-Tutak, F. (2014). Mathematics content knowledge for teaching elementary mathematics: A focus on geometry and measurement. *The Mathematics Enthusiast, 11*(2), 332–338. https://scholarworks.umt.edu/tme/vol11/iss2/7

Bruns, B., & Luque, J. (2015). *Great teachers: How to raise student learning in Latin America and the Caribbean.* World Bank, Washington, DC.

Bryan, P. E. (2014). *A history of the Caribbean Examinations Council, 1973–2013: Regional relevance, international credibility.* North Charleston, North Carolina.

Burnette, D., II. (2019). Just how much does money really matter?: Fresh research bolsters the case for K-12 cashâ€"and a rough road without it. *Education Week, 38*(34), 13. Retrieved from https://www.edweek.org/ew/articles/2019/06/05/student-outcomes-does-more-money-really-matter.html

Caceres, L., de la Peña, J., Di Prisco, C., Pineda, A., & Solotar, A. (2014). Mathematics in Latin American and the Caribbean: Challenges and opportunities. In *INTED2014 Proceedings* (pp. 5366-5366). IATED.

Caemmerer, J. M., Maddocks, D. L. S., Keith, T. Z., & Reynolds, M. R. (2018). Effects of cognitive abilities on child and youth academic achievement: Evidence from the WISC-V and WIAT-III. *Intelligence 68, 6–20.* doi.org/10.1016/j.intell.2018.02.005

Callingham, R., & Watson, J. M. (2017). The development of statistical literacy at school. *Statistics Education Research Journal,16*(1), 181–201. Retrieved from https://iase-web.org/documents/SERJ/SERJ16(1)_Callingham.pdf

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56(2)*, 81–105. doi:10.1037/h0046016

Campos, I. S., Almeida, L. S., Ferreira, A. I., Martinez, L. F., & Ramalho, G. (2013). Cognitive processes and math performance: A study with children at third grade of basic education. *European Journal of Psychology of Education, 28*, 421–436. doi:10.1007/s10212-012-0121-x

Capaldi, M. B. (2019). Mathematics versus statistics. *Journal of Humanistic Mathematic, 9*(2), 149–1 56.

Caribbean Examinations Council. (2017). *Annual Report*. Published by the Caribbean Examinations Council.www.cxc.org

Caribbean Examinations Council. (2018). FAIM data: Electronic processing system.

Caribbean Examinations Council. (n.d). *Procedural guidelines for the development of examinations.*

Caribbean Secondary Education Certificate. (2008). *Mathematics Syllabus.* Effective for examination from May/June 2010. CXC 05/G/SYLL 08. Caribbean Examinations Council. St. Michael: Barbados.

Carlson, M. D. A., & Morrison, R. S. (2009). Study design, precision, and validity in observational studies. *Journal of Palliative Medicine, 12*(1), 77–82. doi:10.1089/jpm.2008.9690

Cass, M., Cates, D., Smith, M., & Jackson, C. (2003). Effects of manipulative instruction on solving area and perimeter problems by students with learning disabilities. *Learning Disabilities, 18*(2), 112–120.

Castle, K., & Needham, J. (2007). First graders' understanding of measurement. Early *Childhood Education Psychology, 35*, 215–221. doi:10.1007/s10643-007-02107

Chan, Y. H. (2003). Basic statistics for doctors. *Singapore Medical Journal, 44*(6), 280–285.

Chester, M. D. (2005a). Making valid and consistent references about school effectiveness from multiple measures. *Educational Measurement: Issues and Practice,24*, 40–50.

Chester, M. D. (2005b). Measuring the impact of state accountability programs. *Educational Measurement: Issues and Practice, 24*, 3–4.

Choi, A., & Jerrim, J. (2016). The use (and misuse) of PISA in guiding policy reform: The case of Spain *Comparative Education*, *52*(2), 230–245. doi.org/10.1080/00131881.2016.1167485

Cirino, P. T., Tolar, T. D., Fuchs, L. S., & Huston-Warren, E. (2016). Cognitive and numerosity predictors of mathematical skills in middle school. *Journal of Experimental Child Psychology 145*, 95–119 .doi.org/10.1016/j.jecp.2015.12.010

Cizek, G., van der Linden, W., & Cook, L. (2012). Testing and data integrity in the administration of statewide student assessment programs. *National council on measurement in education*, 1–12. Retrieved from https://files.eric.ed.gov/fulltext/ED605011.pdf

Cohen, J. (2018). Practices that cross disciplines?: Revisiting explicit instruction in elementary mathematics and English language arts. *Teaching and Teacher Education, 69*, 324–335. doi:10.1016/j.tate.2017.10.021

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston, MA: Houghton Mifflin.

Copp, D. T. (2017). Policy incentives in Canadian large-scale assessment: How policy levers influence teacher decisions about instructional change. *Education Policy Analysis*, *25*(115), 1–26. doi:10.14507/epaa.25.3299

Cormier, D. C., Bulut, O., McGrew, K. S., & Singh, D. (2017). Exploring the relations between Cattell-Horn-Carrol (CHC) cognitive abilities and mathematics achievement. *Applied Cognitive Psychology, 31*, 530-538. doi:10.1002/acp.3350

Cowan, R., Hurry, J., & Midouhas, E. (2018). The relationship between learning mathematics and general cognitive ability in primary school. *British Journal of Developmental Psychology, 36*, 277–284. doi:10.1111/bjdp.12200

Cox, C., & Meckes, L. (2016). International large-scale assessment studies and educational policy making in Chile: contexts and dimensions of influence. *Research Papers in Education, 31*(5), 502–515. doi:10.1080/02671522.2016.122534

Cragg, L., Richardson, S., Hubber, P. J., Keeble, S., & Gilmore, C. (2017). When is working memory important in arithmetic? The impact of strategy and age. *PLos One*, *12*(12), 1–8. doi:10.1371/journal.pone.0188693

Creswell, J. W. (2015). *Educational Research*: *Planning conducting, and evaluating quantitative and qualitative research* (5th ed.). New York: Pearson.

Crompton, H., Grant, M. R., & Shraim, K. Y. H. (2018). Technologies to enhance and extend children's understanding of geometry: A configurative thematic synthesis of the literature. *Journal of Educational Technology & Society, 21*(1), 59–69. Retrieved from https://digitalcommons.odu.edu/teachinglearning_fac_pubs/45

Cullinane, A., & Liston, M. (2016). Review of the leaving certificate biology examination papers (1999–2008) using Bloom's taxonomy–an investigation of the cognitive demands of the examination. *Irish Educational Studies, 35*(3), 249–267. doi:10.1080/03323315.2016.1192480

Cumberbatch, G. (2016, August 12). Decline in cxc math regionally, says examinations council. Retrieved from http://jamaica-gleaner.com/article/news/20160812/decline-cxc-maths-regionally-says-examinations-council

Cunningham, C., Cunningham, S. A., Halim, N., & Yount, K. M. (2019). Public investments in education and children's academic achievements. *Journal of Development Studies*, *55*(11), 2365–2381. doi:10.1080/00220388.2018.1516869

Damiani, V. (2016). Large-scale assessments and educational policies in Italy. *Research Papers in Education, 31*(5), 529–541. doi:10.1080/02671522.2016.1225354

Darwazeh, A. N. (2017). A new revision of the [revised] Bloom's taxonomy. *Distance Learning,14*(3), 13–28. Retrieved from https://search.proquest.com/openview/1947e0c304690943acafde4c4c733ca5/1?pq-origsite=gscholar&cbl=29704

Decker, M., & Bolt, E. (2008). Challenges and opportunities for promoting student achievement through large-scale assessment results, research, reflections and future directions. *Assessment for Effective Intervention, 34*(1), 43–51. doi:10.1177/1534508408314173

di Gropello, E. (2017). *Monitoring educational performance in the Caribbean* (World Bank Paper No.6) Retrieved from https://www.wds.worldbank.org/servlet/WDScontentserver/WDSP/IB/

Doabler, C. T., Clarke, B., Kosty, D., Turtura, J. E., Firestone, A. R., Smolkowski, K., … & Maddox, S. A. (2019). Efficacy of a first-grade mathematics intervention on measurement and data analysis. Exceptional Children, 86(1), 77–94. doi:10.1177/00144029185793

Dobie, T. E. (2019). Expanding conceptions of utility: middle school students' perspectives on the usefulness of mathematics, *Mathematical Thinking and Learning, 21*(1), 28–53. doi:10.1080/10986065.2019.1564969

Donati, G., Meaburn, E., & Dumontheil, I. (2019). The specificity of associations between cognition and attainment in English, math and science during adolescence. *Learning and Individual Differences, 69*, 84–93. doi:10.1016/j.lindif.2018.11.012

Education Act. (2006). Government of St. Vincent and the Grenadines, Ministry of Education, *National Reconciliation, Ecclesiastical Affairs and Information*. Act No. 34. Retrieved from http://education.gov.vc

Education Review Office. (2018). *Keeping children engaged and achieving in mathematics*. New Zealand: Education Review Office Retrieved from http://www.ero.govt.nz/assets/Uploads/ERO-17763-Teaching-Strategies-that-work- Mathematics-v7.pdf.

Ehmke, T., van den Ham, A. K., Sälzer, C., Heine, J., & Prenzel, M. (2020). Measuring mathematics competence in international and national large-scale assessments: Linking PISA and the national educational panel study in Germany. *Studies in*

Educational Evaluation, *65*, 1–10. doi:10.1016/j.stueduc.2020.100847

Elliott, J., Stankov, L., Lee, J., & Beckmann, J. F. (2019). What did PISA and TIMSS ever do for us?: The potential of large-scale datasets for understanding and improving educational practice. *Comparative Education, 55*(1), 133–155. doi:10.1080/03050068.2018.1545386

Ellis, A., Tasova, H. I., & Singleton, B. (2018). How quantitative reasoning can support graph understanding in algebra. In T. E. Hodges, G. I. Roy, & A. M. Tyminski (Eds.), *Proceedings of the 40th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Greenville, SC: University of South Carolina & Clemson University.

Erden, M., & Akgul, S. (2010). Predictive power of math anxiety and perceived social support from teacher for primary students' mathematics achievement. *Journal of Theory and Practice*, *12*(4), 2596–2598

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2013). *G*Power Version 3.1.7* [computer software]. Uiversität Kiel, Germany. Retrieved from http://www.psycho.uni-duesseldorf.de/abteilungen/aap/gpower3/download-and-register

Ferrini-Mundy, J. (2000). Principles and standards for school mathematics: A guide for mathematicians. *The American Mathematical Society, 47*(8), 868–876. Retrieved from https://www.aps.edu/academics/common-core-state-standards/documents/comm-ferrini.pdf

Floyd, R. G., Evans, J. J., & McGrew, K. S. (2003). Relationships between measures of

Cattell-Horn-Carrol (CHC) cognitive abilities and mathematics achievement across the school age years. *Psychology in Schools, 40*(2), 155–171. doi:10.1002/pits.10083.

Forehand, M. (2010). Bloom's taxonomy: Emerging perspectives on learning. In L. Dunegan. *Teaching and Technology* (pp. 1–10.). Retrieved from http://epltt.coe.uga.edu/index.php?title=Bloom%27s_Taxonomy.

Fraiser, A. (2019, October 25). *1979*–The year of independence. *Searchlight*. https://searchlight.vc/searchlight/dr-fraser/2019/10/25/1979-the-year-of-independence-continued-from-last-week/

Frankfort-Nachmias, C., Nachmias, D., & DeWaard, J. (2015). *Research methods in the social sciences* (8th ed.). New York, NY: Worth.

Galitskaya, V. N., & Drigas, A. S. (2020). Special education: Teaching geometry with ICTs. *International Journal of Emerging Technologies in Learning, 15*(6), 17 – 182. doi:10.3991/ijet.v15i06.11242

García-Pérez, M. A. (2012). Statistical conclusion validity: Some common threats and simple remedies. *Quantitative Psychology and Measurement, 3*(325), 1–11. doi:10.3389/fpsyg.2012.00325

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology, 47*(6), 1539–1552. doi:10.1037/a0025510

Gelbart, D. (2007). *Cognitive abilities that underlie mathematics achievement: A high ability perspective* (T). University of British Columbia. Retrieved from https://open.library.ubc.ca/collections/ubctheses/831/items/1.0054546

George, A., & Robitzsch, A. (2018). Focusing on interaction between content and cognition: A new perspective on gender differences in mathematical sub-competencies. *Applied Measurement in Education, 31*(1), 79–97. doi:10.1080/08957347.2017.1391260

Gierl, M. J. (1997). Comparing cognitive representations of test developers and students on a mathematics test with Bloom's taxonomy, *The Journal of Educational Research,91(*1), 26–32. doi:10.1080/00220679709597517

Gilmore, C., Clayton, S., Cragg, L., McKeaveney, C., Simms, V., & Johnson, S. (2018). Understanding arithmetic concepts: The role of domain-specific and domain-general skills. *PLoS ONE, 19*(3), 1–20. doi:10.1371/journal.pone.0201724

Gimbert, F., Camos, V., Gentaz, E., & Mazens, K. (2019). What predicts mathematics achievement? Developmental changes in 5– and 7– year -old children. *Journal of Experimental Child Psychology, 178*, 104–120. doi:10.1016/j.jecp.2018.09.013

Giofrè, D., Borella, E., & Mammarella, I. C. (2017). The relationship between intelligence, working memory, academic self-esteem, and academic achievement. *Journal of Cognitive Psychology, 29*(6), 731–747. doi:10.1080/20445911.2017.1310110

Giofrè, D., Mammarella, I. C., & Cornoldi, C. (2014). The relationship among geometry, working memory, and intelligence in children. *Journal of Experimental Child*

Psychology, *123*, 112–128. doi:10.1016/j.jecp.2014.01.002

Giofrè, D., Mammarella, I. C., Ronconi, L., & Cornoldi, C. (2013). Visuospatial working memory in intuitive geometry, and in academic achievement in geometry. *Learning and Individual Differences, 23*, 114–122. doi:10.1016/j.lindif.2012.09.012

Goertz, M., & Duffy, M. (2003). Mapping the landscape of high-stakes testing and accountability programs. *Theory Into Practice, 42*(1), 4–8. doi:10.1207/s15430421tip4201_2

Goldstein, H. (2007). The future of statistics within the curriculum. *Teaching Statistics, 29*(1), 8–9. doi:10.1111/j.1467-9639.2006.00259.x

Granello, D. H. (2000). Encouraging the cognitive development of supervisees: Using Bloom's taxonomy in supervision. *Counselor Education & Supervision, 40*(1), 31–46. doi:10.1002/j.1556-6978.2000.tb01797.x

Granello, D. H. (2001). Promoting cognitive complexity in graduate written work: Using Bloom's taxonomy as a pedagogical tool to improve literature reviews. *Counselor Education & Supervision, 40*(4) 292–307. doi:10.1002/j.1556-6978.2001.tb01261.x

Green, C. T., Bunge, S. A., Chiongbian, V. B., Barrow, M., & Ferrer, E. (2017). Fluid reasoning predicts future mathematical performance among children and adolescents. *Journal of Experimental Child Psychology, 157*, 125–143. doi:10.1016/j.jecp.2016.12.005

Greeno, J. G., & Collins, A. (2008). Commentary on the final report of the national mathematics advisory panel. *Educational Researcher*, *37*, 618– 623.

doi:10.3102/0013189X08327997

Grek, S. (2013). Expert moves: International comparative testing and the rise of expertocracy. *Journal of Education Policy, 28*(5), 695–709. doi:10.1080/02680939.2012.758825

Gustafsson, J. E. (2008). Effects of international comparative studies on educational quality on the quality of educational research. *European Educational Research Journal, 7*(1), 1–17. doi:10.2304/eerj.2008.7.1.1.

Güvendir, M. A. (2017). Determination of the relationship between the students' mathematical literacy and home and school educational resources in program for international student assessment–(PISA 2012). *Mersin University Journal of the Faculty of Education, 13* (1), 94–109. doi:10.17860/mersinefd.305762

Hadzhikoleva, S., Hadzhikolev, E., & Kasakliev, N. (2019). Using peer assessment to enhance higher order thinking skills. *Technology Education Management Informatics, 8*(1), 242–247. doi:10.18421/TEM81-34

Hartung, D. M., & Touchette, D. (2009). Overview of clinical research design. *American Journal of Health-System Pharmacy*, *66*(4) 398–408. doi:10.2146/ajhp080300

Hassan, M. N., Abdullah, A. H., Ismail, S., Suhud, A., & Hamzah, M. H. (2019). Mathematics curriculum framework for early childhood education based on science, technology, engineering and mathematics (STEM). *International Electronic Journal of Mathematics Education, 14*(1), 15–31. doi:10.12973/iejme/3960

Heyneman, S. P., & Lee, B. (2013). Impact of international studies of academic achievement on policy and research. In L. Rutkowski, M. von Davier & D. Rutkowski, (Eds.). *Handbook of international large-scale assessment*. New York: Taylor & Francis Group. doi:10.1201/b16061

Hosein, A. (2019). Girls' video gaming behaviour and undergraduate degree selection: A secondary data analysis approach. *Computers in Human Behavior, 91*, 226–235. doi:10.1016/j.chb.2018.10.001

Howie, S. J., & Plomp, T. (2006). Raising achievement levels in mathematics and science internationally: What really matters. In S. J. Howie & T. Plomp (Eds.) *Contexts of learning mathematics and science. Lessons learned from TIMSS*. (pp. 71–82). London: Routledge.

Hudson, J. I., Pope, H. G., Jr., & Glynn, R. (2005). The cross-sectional cohort study: An underutilized design. *Epidemiology, 16*(3), 355–359. doi:10.1097/01.ede.0000158224.50593.e3

Hurrell, D. (2015). Measurement: Five considerations to add even more impact to your program. *Australian Primary Mathematics Classroom, 20*(4), 14–18. Retrieved from https://researchonline.nd.edu.au/edu_article/170

Husén, T. (1979). An international research venture in retrospect: The IEA surveys. *Comparative Education Review, 23*(3), 371–385. doi:10.1086/446067

İnce, M., & Gözütok, F. D. (2018). Effect of parental education and home educational resources to students' results of PISA reading skills test. *İlköğretim Online, 17*(2), 947–958. doi:10.17051/ilkonline.2018.419346

Irvine, J. (2017). A comparison of revised Bloom and Marzano's new taxonomy of learning. *Research in Higher Education Journal*, 33, 1–16 Retrieved from https://eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=EJ1161486

Ito, P. K. (1980). 7 robustness of ANOVA and MANOVA test procedures. In P. R. Krishnaiah (Ed.), *Handbook of statistics, 1* 199–236. doi:10.1016/S0169-7161(80)01009-7

Jackson, J. F. L., O'Callaghan, E. M., & Adserias, R. P. (2014). Approximating glass ceiling effects using cross-sectional data. *New Directions for Institutional Research 159*. doi:10.1002/ir

Jiang, R., Li, X., Xu, P., & Chen, Y. (2019). Inhibiting intuitive rules in a geometry comparison task: Do age level and math achievement matter? *Journal of Experimental Child Psychology, 186,* 1–16. doi:10.1016/j.jecp.2019.05.003

Johansson, S. (2016). International large-scale assessments: what uses, what consequences? *Educational Research, 58*(2), 139–48. doi:10.1080/00131881.2016.1165559

Jordan, N. C., Glutting, J., & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and Individual Difference, 20*, 82–88. doi:10.1016/j.lindif.2009.07.004

Jules, D. (2012). *Math coordinator set to address low cxc math pass rate*. Retrieved from https://antiguaobserver.com/math-co-ordinator-set-to-address-low-cxc-math-pass-rate/

Karakolidis, A., Pitsia, V., & Emvalotis, A. (2016). Examining students' achievement in mathematics: A multilevel analysis of the program for international student assessment (PISA) 2012 data for Greece. *International Journal of Educational Research, 79*, 106–115. doi:10.1016/j.ijer.2016.05.013

Khan, S. A. (2017). *Mathematics proficiency of primary school students in Trinidad and Tobago* (Doctoral Dissertation). Retrieved from https://www.researchgate.net/publication/328433370

Kieran, C. (2013). The false dichotomy in mathematics education between conceptual understanding and procedural skills: An example from algebra. In K. R. Leatham (Ed.), *Vital directions for mathematics education research* (pp. 153–171). New York: Springer.

Kilicoglu, E. (2020). A Study on the professional development of prospective primary school teachers: subject matter knowledge on geometry, *International Online Journal of Educational Sciences, 12* (2), 30–54. doi:10.15345/iojes.2020.02.003

Klieger, A. (2016). Principals and teachers: Different perceptions of large-scale assessment. *Journal of Educational Research, 75*, 134–145. doi:10.1016/j.ijer2015.11.006

Klinger, D., DeLucas, C., & Miller, T. (2008). The evolving culture of large-scale assessment in Canadian education. *Canadian Journal of Educational Administration and Policy, 76*, 1–34.

Knuth, E., Stephens, S., Blanton, M., & Gardiner, A. (2016). Build an early foundation for algebra success. *Kappan,79*(6). kappanmagazine.org

Konietschke, F., Bathke, A., Harrar, S., & Pauly, M. (2015). Para-metric and nonparametric bootstrap methods for general MANOVA. *Journal of Multivariate Analysis*, *140*, 291–301. doi:10.1016/j.jmva.2015.05.001

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*(4), 212–218.

Lake, D., & Kemp, N. (2001). Choosing and using graphs. *Australian Mathematics Teacher, 57*(3), 7–12. Retrieved from https://aamt.edu.au/Journals/Journals-Index/The-Australian-Mathematics-Teacher

Larson, M. J., & Whitin, D. J. (2010, Fall). Young children use graphs to build mathematical reasoning. *Dimensions of Early Childhood, 38*(3), 16–21. Retrieved from https://eric.ed.gov/?id=EJ945684

Leacock, C. (2015). Status of mathematics education in the Eastern Caribbean: Issues and possible solutions for teacher preparation and support. In D. Marshall, L. Lazarus, J. Smith, & J. Thompson (Eds.), *Journal of Eastern Caribbean Studies*, *40*(3), 210–233. Retrieved from https://www.jecsonline.com/wp-content/uploads/2019/09/JECS-Vol-40-No.-3-Dec-2015.pdf#page=214

Lee, J., Collins, D., & Melton, J. (2016). What does algebra look like in early childhood? *Childhood Education, 92*(4), 305–310. doi:10.1080/00094056.2016.1208009

Lee, J., & Stankov, L. (2018). Non-cognitive predictors of academic achievement: Evidence from TIMSS and PISA. *Learning and Individual Differences, 65*, 50–64. doi:10.1016/j.lindif.2018.05.009

Lee, K., & Bull, R. (2016). Developmental changes in working memory, updating, and math achievement. *Journal of Educational Psychology, 108*(6), 869–882. doi:10.1037/edu0000090

Lee, K., Ng, S. F., & Bull, R. (2018). Learning and solving algebra word problems. The role of relational skills, arithmetic, and executive functioning. *Developmental Psychology, 54*(9), 1758–1772. doi:10.1037/dev0000561

Levin, K. A. (2006). Study design III: Cross-sectional studies. *Evidence-Based Dentistry, 7*(1), 24–25. doi:10.1038/sj.ebd.6400375

Lewis, S. (2017). Communities of practice and PISA for schools: Comparative learning or a mode of educational governance? *Education Policy Analysis Archives*, *25*(92), 1–21. doi:10.14507/epaa.25.2901

Lindquist, M., Philpot, R., Mullis, I. V. S., & Cotter, K. E. (2019). TIMSS 2019 mathematics framework. In I. V. S. Mullis & M. O. Martin (Eds.), *TIMSS 2019 assessment frameworks* (pp. 11–25). Chestnut Hill, MA: TIMSS & PIRLS International Study Center.

Lipnevich, A. A., Preckel, F., & Krumm, S. (2016). Mathematics attitudes and their unique contribution to achievement: Going over and above cognitive ability and personality. *Learning and Individual Differences, 47*, 70–79. doi:10.1016/j.lindif.2015.12.027

Lipscomb, J. W., Jr. (1985). Is Bloom's taxonomy better than intuitive judgement for classifying test questions. *Education, 106*(1), 102–107.

Litke, E. (2020a). Instructional practice in algebra: Building from existing practices to inform an incremental improvement approach. *Teaching and Teacher Education, 91*, 1–12. doi:10.1016/j.tate.2020.103030

Litke, E. (2020b). The nature and quality of algebra instruction: Using a content-focused observation tool as a lens for understanding and improving instructional practice. *Cognition and Instruction*, *38*(1), 57–86. doi:10.1080/07370008.2019.1616740

Litke, E., & Corven, J. (2019). Examining features of algebra instruction in classrooms using a reform-oriented curriculum. In S. Otten, A. G. Candela, Z, de Araujo, C. Haines, & C. Munter (Eds.), *Proceedings of the forty-first annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. St Louis, MO: University of Missouri.

Locuniak, M. N., & Jordan, N. C. (2008). Using kindergarten number sense to predict calculation fluency in second grade. *Journal of Learning Disabilities*, *4*(5), 451–459. doi:10.1177%2F0022219408321126

Looney, J. W. (2011, April). *Integrating formative and summative assessment: Progress toward a seamless system?* OECD Education Working Papers, No. 58, (pp. 1–64). OECD Publishing. doi:10.1787/5kghx3kbl734-en

MacGregor, M. (2004). Goals and content of an algebra curriculum for the compulsory years of schooling. In K. Stacey, H. Chick & M. Kendal (eds.), *The future of the teaching and learning of algebra.* The 12th ICMI Study. New ICMI Study Series, (vol.8, pp. 311–328). Dordrecht, the Netherlands: Kluwer Academic. doi:10.1007/1-4020-8131-6_12

Männamaa, M., Kikas, E., Peets, K., & Palu, A. (2012). Cognitive correlates of math skills in third-grade students. *Educational Psychology, 32*(1), 21–44. doi:10.1080/01443410.2011.621713

Martin, M. O., & Mullis, I. V. S. (2019). TIMSS 2015: Illustrating advancements in large-scale international assessments. *Journal of Educational and Behavioral Statistics*, *44* (6), 752–781. doi:10.3102/1076998619882030

Martin, M. O., Mullis, I. V. S., & Hooper, M. (2015). (Eds.). *Methods and procedures in TIMSS 2015*. Boston College: TIMSS and PIRLS international study center: Retrieved from http://timssandpirls.bc.edu/publications/timss/2015-methods.html

Marzano, R. J., & Kendall, J. S. (2006). *The new taxonomy of educational objectives* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Mayr, S., Erdfelder, E., Buchner, A., & Faul, F. (2007). A short tutorial of GPower. *Tutorials in Quantitative Methods for Psychology, 3*(2), 51–59. Retrieved from https://www.psychologie.hhu.de/fileadmin/redaktion/Oeffentliche_Medien/Fakultaeten/Mathematisch-Naturwissenschaftliche_Fakultaet/Psychologie/AAP/Publikationen/2007/Mayr_Erdfelder_Buchner_Faul__2007_.pdf

McGrew, K. S., LaForte, E. M., & Schrank, F. A. (2014). *Technical manual: Woodcock Johnson IV*. Rolling Meadows, IL: Riverside.

Miller, T. (2013). Using large-scale assessment scores to determine student grades. *Canadian Journal of Education, 36*(3), 317–353. Retrieved from https://files.eric.ed.gov/fulltext/EJ1057970.pdf

Mills, J. D. (2004). Students' attitudes toward statistics: Implications for the future. *College Student Journal, 38*(3), 349–361. Retrieved from https://go.gale.com/ps/anonymous?id=GALE%7CA123321892&sid=googleScholar&v=2.1&it=r&linkaccess=abs&issn=01463934&p=AONE&sw=w

Mji, A., & Makgato, M. (2006). Factors associated with high school learners' poor performance: A spotlight on mathematics and physical science. *South African Journal of Education*, *26*(2), 253–266. Retrieved from http://www.sajournalofeducation.co.za/index.php/s aje/article/view/80/55

Monteith, L. (2016). Math performance blamed for dip in exam results. [Web log comment]. Retrieved from https://www.caymancompass.com/2016/10/26/math-performance-blamed-for-dip-in-exam-results/

Moore, D. S. (2007). New pedagogy and a new content: The case of statistics. *International Statistical Review*, *65*(2), 123–165. doi:10.1111/j.1751-5823.1997.tb00390.x

Morsanyi, K., Prado, J., & Richard, E. (2018). Editorial: The role of reasoning in mathematical thinking. *Thinking & Reasoning*, *24*(2), 129–137. doi.org/10.1080/13546783.2018.1435425 Moses, R. P. (2001). *Radical equations: Math literacy and civil rights*. Boston MA: Beacon Press.

Moses, R. P., & Cobb, C. E. (2001). Radical equations: Organizing math literacy in America's schools. Beacon Press. Retrieved from http://ebookcentral.proquest.com/lib/waldenu/detail.action?docID=3117955.

Mullis, I. V. S., & Martin, M. O. (Eds.). (2019). *TIMSS assessment frameworks*. IAE International Study Center: School of Education, Boston College. Retrieved from

Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (Eds.). (2012). *TIMSS 2011 International Results in Mathematics: International study.* TIMSS & PIRLS, Boston College.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2015). *TIMSS* 2015 international results in mathematics. *IEA TIMSS & PIRLS International Study Centre*. Boston College: Lynch School of Education. Retrieved from http://timssandpirls.bc.edu/timss2015/international-results/http://TIMSSandpirls.bc.edu/TIMSS2011/downloads/T11_IR_Mathematics_FullBook.pdf

Musso, M. F. B., Boekaerts, M., Segers, M., & Cascallar, E. C. (2019). Individual differences in basic cognitive processes and self-regulated learning: Their interaction effects on math performance. *Learning and Individual Differences, 71*, 58–70. doi:10.1016/j.lindif.2019.03.003

Naidoo, J., & Kapofu, W. (2020). Exploring female learners' perceptions of learning geometry in mathematics. *South African Journal of Education, 40* (1), 1–11. doi:10.15700/saje.v40n1a1727

National Center for Education Statistics. (2016). The condition of education 2016 (NCES 2016–144). *National Center for Education Statistics, Institute of Education Sciences*. Washington, DC: U.S. Department of Education. https://files.eric.ed.gov/fulltext/ED565888.pdf

National Council of Teachers of Mathematics. [NCTM]. (2006). *Curriculum focal points for prekindergarten through grade 8 mathematics: A quest for coherence*. Reston, VA: Author.

Nimon, K. F. (2012). Statistical assumptions of substantive analyses across the general linear model: A mini-review. *Frontiers in Psychology, 3*(322), 1–5. doi: 10.3389/fpsyg.2012.00322

Ninomiya, S. (2019). The impact of PISA and the interrelation and development assessment policy and assessment theory in Japan. *Assessment in Education: Principles, Policy & Practice, 26*(1), 91-110. doi:10.1080/0969594X.2016.1261795

Nitko, A. J. (2004). *Educational assessment of students* (4th ed.). Upper Saddle River, New Jersey: Pearson Merrill Prentice Hall.

O'Connell, M. (2018). The power of cognitive ability in explaining educational test performance, relative to other ostensible contenders. *Intelligence, 66,* 122–1127. doi:10.1016/j.intell.2017.11.011

O'Connor, K. (2017). *How to grade for learning, linking grades to standards* (4th ed.). Thousand Oaks, CA: Corwin.

OECD. (2016). *FDI in figures*. Retrieved from www.oecd.org/corporate/FDI-In-Figures-April-2016.pdf

Ölmez, I. B. (2020). Modelling mathematics achievement in hierarchical linear models. *Elementary Education Online, 19* (2), 944–957. doi:10.17051/ilkonline.2020.695837

Olson, C. L. (1976). On choosing a test statistics in multivariate analysis of variance. *Psychological Bulletin*, 579–586. doi:10.1037/0033-2909.83.4.579

Organization for Economic Co-operation and Development (OECD). (2004). *Learning for tomorrow's world: First results from PISA 2003*. Paris: OECD.

Organization for Economic Co-operation and Development (OECD). (2014). PISA 2012 Results: What Students Know and Can Do (Volume I, Revised edition, February 2014): *Student Performance in Mathematics, Reading and Science*, Paris: OECD. doi:10.1787/9789264208780-6-en

Osborne, J. W. (2010). Improving your data transformations: Applying the Box-Cox transformation. *Practical Assessment, Research, and Evaluation, 15*(12), 1–9. doi:10.7275/qbpc-gk1

Ozga, J. (2012). Governing knowledge: Data, inspection and education policy in Europe. *Globalisation, Societies and Education, 10*(4), 439–455. doi:10.1080/14767724.2012.735148

Panizzon, D., Geer, R., Paige, K., O'Keeffe, L., Schultz, L., Zeegers, Y., & Brown, L. (2018). Exploring the 'hard facts' around STEM in Australia: Females, low socio-economic status and absenteeism. [Special Issue: STEM Education]. *International Journal of Innovation in Science and Mathematics Education, 26*(8), 30–44. Retrieved from https://openjournals.library.sydney.edu.au/index.php/CAL/article/view/12961/11906

Passolunghi, M. C., Cargnelutti, E., & Pellizzoni, S. (2019). The relation between cognitive and emotional factors and arithmetic problem-solving. *Educational Studies in Mathematics, 100*, 271–290. doi:10.1007/s10649-018-9863-y

Patahuddin, S. M., & Lowrie, T. (2019). Examining teachers' knowledge of line graph task: A case of travel task. *International Journal of Mathematics Education, 19*, 781–800. doi:10.1007/s10763-018-9893-z

Paules, C. I., Marston, H. D., & Fauci, A. S. (2019). Measles in 2019: Going backward. *New England Journal of Medicine*. doi:10.1056/NEJM p1905099?articleTools=true

Pedersen, I. F. (2015). What characterizes the algebraic competence of Norwegian upper secondary school students? Evidence from TIMSS advanced. *International Journal of Science and Mathematics Education, 13*(1), 73–95. doi:10.1007/s10763-013-9468-y

Peng, P., Namkung, J., Barnes, M., & Sun, C. (2016). A meta-analysis of mathematics and working memory: Moderating effects of working memory domain, type of mathematics skill, and sample characteristics. *Journal of Educational Psychology, 108*(4), 455–473. doi:10.1037/edu0000079.

Peng, P., Wang, C., Wang, T., & Lin, X. (2019). A meta-analysis of the relation between fluid intelligence and reading/mathematics: Effects of tasks, age, and social economics status. *Psychological Bulletin, 145*(2), 189–236. doi:10.1037/bul0000182

Phelps, R. (2012). The effects of testing on achievement; Meta-analysis and research summary, 1910 – 1920. *International Journal of Testing, 12*, 21–43. doi:10.1080/15305058.2011.602920

Pituch, K. A., & Stevens, J. P. (2016). *Applied multivariate statistics for the social sciences* (6th ed.). New York: Routledge.

Podsakoff, P. M., MacKenzie, S. B., Lee, J., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*, 897–903. doi:10.1037/0021-9010.88.5.879.

Popham, J. W. (2000). *Educational measurement: Practical guidelines for educational leaders* (3rd ed.). Boston: Allyn and Bacon.

Popham, J. W. (2002). *Classroom assessment: What teachers need to know* (3rd ed.). Boston: Allyn and Bacon.

Prendergast, M., & Treacy, P. (2018). Curriculum reform in Irish secondary schools a focus on algebra. *Journal of Curriculum Studies*, *50*(1), 126–143. doi:10.1080/00220272.2017.1313315

Price, J. H., & Murnan, J. (2013). Research limitations and the necessity of reporting them. *American Journal of Health Education*, *35*(2), 66–67. doi:10.1080/19325037.2004.10603611

Primi, C., Bacherini, A., Beccari, C., & Donati, M. A. (2020). Assessing math attitude through the attitude toward mathematics – short form in introductory statistics course students. *Studies in Educational Evaluation*, *64*, 1–9. doi:10.1016/j.stueduc.2020.100838

Primi, R., Ferrão, M. E., & Almeida, L. S. (2010). Fluid intelligence as a predictor of learning: A longitudinal multilevel approach applied to math. *Learning and Individual Differences, 20,* 446–451. doi:10.1016/j.lindif.2010.05.001

Prince, S. C. (2018). *Education minister says education revolution responsible for change in mind set*. Retrieved from http://education.gov.vc/education/index.php/news/555-education-minister-says-education-revolution-responsible-for-change-of-mindset-in-svg

Psacharopoulos, G., & Patrinos, H. A. (2018). Returns to investment in education: A decennial review of the global literature. *Education Economics, 26*(5), 445–458. doi:10.1080/09645292.2018.1484426

Quinn-Leandro, J. (2011). *Ministry of education working to double CXC passes in maths.* Retrieved from http://www.antigua.gov.ag/article_details.php?id=2057&category=38

Quinn-Leandro, J. (2012). *Math stats not adding up*. Retrieved from https://barbadostoday.bb/2012/08/19/math-stats-not-adding-up/

Radhakrishnan, G. (2013). Non-experimental research designs: Amenable to nursing contexts. *Asian Journal of Nursing Education and Research, 3*(1). 25–28. Retrieved from https://ajner.com/

Ramirez, T. V. (2017). On pedagogy of personality assessment: Application of Bloom's taxonomy of educational objectives. *Journal of Personality Assessment, 99*(2), 146–152. doi:10.1080/00223891.2016.1167059

Rau, M. A., & Mathews, P. G. (2017). How to make 'more' better? Principles for effective use of multiple representations to enhance students' learning about fractions. *ZDM Mathematics Education*, *49*(4),531–544. doi:10.1007/s11858-017-0846-8

Rautalin, M., & Alasuutari, P. (2009). The uses of the national PISA results by Finnish officials in central government. *Journal of Education Policy, 24*(5), 539–556. doi:10.1080/02680930903131267

Ray, J. V. (2020). *Cross-sectional research designs in criminology and criminal justice*. New York: Oxford University Press.

Reid, R. (2011). *Reid examines Jamaica's poor cxc results*. Retrieved from http://jamaica-gleaner.com/gleaner/20110223/news/news1.html

Reio, T. G., Jr. (2010). The threat of common method variance bias to theory building. *Human Resource Development Review, 94*(4). 405 – 411. doi:10.1177/1534484310380331

Reyna, V. F., & Brainerd, C. J. (2007). The importance of mathematics in health and human judgement: Numeracy, risk communication, and medical decision making. *Learning and Individual Differences, 17*, 147–159.

Reys, R. E., Lindquist, M. M., Lambdin, D. V., Smith, N. L., Rogers, A., Falle, J., … & Bennett, S. (2012). *Helping children learn mathematics.* (1st Australian Ed.). Milton, Qld: John Wiley & Sons, Australia.

Rindfleisch, A., Malter, A. J., Ganesan, S., & Moorman, C. (2008). Cross-sectional versus longitudinal survey research: Concepts, findings, and guidelines. *Journal of Marketing Research*, *45*(3), 261–279.

Roderick, M., & Engel, M. (2001). The grasshopper and the ant: Motivational responses of low achieving students to high-stakes testing. *Educational Analysis and Policy Analysis*, *23*(3), 197–227.

Roegner, K. (2013). *Cognitive levels and approaches taken by students failing written examinations in mathematics*. Department of Mathematics, Technical University Berlin, Germany.

Rogers, W. T. (2014). Improving the utility of large-scale assessment in Canada. *Canadian Journal Education*, *37*(3), 1–22. Retrieved from https://files.eric.ed.gov/fulltext/EJ1057942.pdf

Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnic, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence*, *53*, 118–137. doi:10.1016/j.intell.2015.09.002

Rutkowski, D., & Delandshere, G. (2016). Causal inferences with large scale assessment data; using a validity framework. *Large Scale Assessment in Education*, *4*(6), 2–18. doi:10.1186/s40536-016-0019-1

Sahlberg, P. (2011). PISA in Finland: an education miracle or an obstacle to change? *Center for Educational Policy Studies Journal, 1*(3), 119–140. Retrieved from https://files.eric.ed.gov/fulltext/EJ1130792.pdf

Sakia, R. M. (1992). The Box-Cox transformation technique: A review. *The statistician, 41,* 169–178. doi:10.3102%2F0013189X031001013

Schoenfeld, A. (1995). Is thinking about algebra a misdirection? In C. Lacampagne, W. Blair, & J. Kaput (Eds.), *The algebra colloquium,2: Working group papers* (pp. 83–86). Washington, DC: U.S. Department of Education.

Schoenfeld, A. H. (2002). Making mathematics work for all children: Issue of standards, testing and equity. *Educational Researcher*, *31*(1), 13–25. doi:10.3102/0013189X031001013

Seitz, P. (2017). Curriculum alignment among intended, enacted and assessed curricula for grade 9 mathematics. *Journal of Canadian Studies*, *15*(1), 72–9 4. Retrieved from https://pdfs.semanticscholar.org/16ae/5e176e891456f3fb489cdde1483449dcd0cd.pdf

Sellar, S., & Lingard, B. (2014). The OECD and the expansion of PISA: New global modes of governance in education. *British Educational Research Journal 40* (6), 917–936. doi:10.1002/berj.3120

Semeraro, C., Giofrè, D., Coppola, G., Lucangeli, D., & Cassibba, R. (2020). The role of cognitive and non-cognitive factors in mathematics achievement: The importance of the quality of the student-teacher relationship in middle school. *PLoS ONE 15*(4): doi:10.1371/journal.pone.0231381

Serow, P., Callingham, R., & Muir, T. (2014). *Primary mathematics: Capitalising on ICT for today and tomorrow.* Melbourne: Cambridge University Press.

Shephard, L. A. (2010). What the marketplace has brought us: Item-by-item teaching with little instructional insight. *Peabody Journal of Education, 85*, 246–257. doi:10.1080/01619561003708445

Skagerlund, K., & Träff, U. (2016). Processing of space, time, and number contributes to mathematical abilities above and beyond domain-general cognitive abilities. *Journal of Experimental Child Psychology 143*, 85–101. doi:10.1016/j.jecp.2015.10.016

Sodha, J. (2012). *Solving the math problem*. Retrieved from https://barbadostoday.bb/2012/11/07/solving-the-maths-problem/

Soni, A., & Kumari, S. (2017). The role of parental math anxiety and math attitude in the children math achievement. *International Journal of Science and Mathematics Education, 15*, 331–347. doi:10.1007/s10763-015-9687-5

Spector, P. E. (2019). Do not cross me: Optimizing the use of cross-sectional designs. *Journal of Business and Psychology, 34*, 125–137. doi:10.1007/s10869-018-09613-8

Spiegelhalter, D. (2020). The art of statistics: How we learn from data. *Numeracy, 13*, 1–4. doi:10.5038/1936-4660.13.1.7

Squires, D. (2012). Curriculum alignment research suggests that alignment can improve student achievement. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *85*(4),129–135. doi:10.1080/00098655.2012.657723

Star, J. R., Caronongan, P., Foegen, A., Furgeson, J., Keating, B., Larson, M. R., ... Zbiek, R. M. (2015). *Teaching strategies for improving algebra knowledge in middle and high school students* (NCEE 2014-4333).

STEM Education Coalition. (2009). Retrieved from http://www.stemedcoalition.org/content/objectives/Default.aspx

Stylianou, D. A., Stroud, R., Cassidy, M., Knuth, E., Stephens, A., Gardiner, A., & Demers, L. (2019). Putting early algebra in the hands of elementary school teachers: Examining fidelity of implementation and its relation to student performance. *Journal for the Study of Education and Development*, *42*(3), 523–569. doi:10.1080/02103702.2019.1604021

Sui Chu Ho, E. (2016). The use of large-scale assessment (PISA): Insights for policy and practice in the case of Hong Kong. *Research Papers in Education, 31*(5), 516–528. doi:10.1080/02671522.2016.1225351

Syarifudin, S., Purwanto, P., Irawan, E. B., Sulandra, I. M., & Fikriyah, U. (2019). Student verbal interaction in geometry problem-solving through cognitive activities. *International Journal of Instruction*, *12*(3), 167–182. doi:10.29333/iji.2019.12311a

Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics* (3rd ed.). New York, NY. Harper Collins.

Tajudin, N. M., & Chinnappan, M. (2016). The link between higher order thinking skills, representation and concepts in enhancing TIMSS tasks. *International Journal of Instruction, 9*(2), 1308–1470. doi:10.12973/iji.2016.9214a

Tarman, B., & Kuran, B. (2015). Examination of cognitive level of questions in social studies textbooks and the views of teachers based on Bloom taxonomy. *Education Sciences: Theory and Practice, 15*(1), 213–222. doi:10.12738/estp.2015.1.2625

Taub, G. E., Floyd, R. G., Keith, T. Z., & McGrew, K. S. (2008). Effects of general and broad cognitive abilities on mathematics abilities. *School Psychology Quarterly, 23*, 187–198. doi:10.1037/1045-3830.23.2.187

Thorndike, R. M. (1997). *Measurement and evaluation in psychology and education* (6th Ed.). New Jersey: Merrill.

Tirza, J. (2020). Constructivism and teaching-learning Process in Christian School: Case Study in a Christian school in Tangerang, Indonesia. *International Journal of Psychosocial Rehabilitation*, *24*(4), 584–590. doi:10.37200/IJPR/V24I4/PR201036

Tolar, T. D., Fuchs, L., Fletcher, J. M., Fuchs, D., & Hamlett, C. L. (2016). Cognitive profiles of mathematical problem-solving learning disability for different definitions of disability. *Journal of Learning Disabilities, 49*(3), 240–256. doi:10.1177/0022219414538520

Trochim, W. M. (2006). *The research methods knowledge base*. (2nd Ed.). Retrieved from http://www.socialresearchmethods.net/kb/

United Nations World Population Prospects. (2019). Revision worldpopulationreview.com. PAHO. Pan American Health Organization about Saint Vincent and the Grenadines.

Ünlü, M., & Ertekin, E. (2017). A structural equation model for factors affecting eighth graders' geometry achievement. *Educational Sciences: Theory & Practice,* 1815–1846. doi:10.12738/estp.2017.5.0545

Ursani, A. A., Memon, A. A., & Chowdhry, B. S. (2014). Bloom's taxonomy as a pedagogical model for signals and systems. *International Journal of Electrical Engineering Education, 51*(2). 162–173. doi:10.7227/IJEEE.51.2.7

Üstün, I., & Ubuz, B. (2004). *Geometrik kavramların geometer's sketchpad yazılımı ile geliştirilmesi* [Development of geometric concepts with geometer's sketchpad software]. Paper presented at the Eğitimde İyi Örnekler Konferansı, Sabancı University, İstanbul, Turkey.

VanDerHeyden, A. M., & Burns, M. K. (2009). Performance indicators in math: Implications for brief experimental analysis of academic performance. *Journal of Behavioral Education, 18*(1), 71–91. Retrieved from https://www.springer.com/journal/10864

Van de Walle, J. A., Karp, K. S., & Bay-Williams, J. M. (2013). *Elementary and middle school mathematics: Teaching developmentally.* (8th ed.). Boston: Pearson.

Vista, A. (2016). The role of PS ability and RC skill in predicting growth trajectories of mathematics achievement. *Cogent Education, 3,* 1–22. doi:10.1080/23331186x.2016.1222720

Wagemaker, H. (2013). International large-scale assessments. In Rutkowski, L., von Davier, M., & Rutkovski, D. (Eds.), *Handbook of International large-scale assessment: Background, Technical Issues, and Methods of Data Analysis* (pp. 11–36). Florence, KY: Chapman & Hall/CRC.

Watson, J. M. (2001). Profiling teachers’ competence and confidence to teach particular mathematics topics: The case of chance and data. *Journal of Mathematics Teacher Education, 4*, 305–337.

Waxman, D. (2020). Is mathematics unreasonably effective? *Australasian Journal of Philosophy*. doi:10.1080/00048402.2020.1724164

Whittle, R. J., Benson, A. C., Ullah, S., & Telford, A. (2018). Investigating the influence of question type and cognitive process on academic performance in VCE physical education: A secondary data analysis. *Educational Research and Evaluation, 24*(8), 504–522. doi:10.1080/13803611.2019.1612256

Wong, T. T., & Ho, C. S. (2017). Component processes in arithmetic word-problem solving and their correlates. *Journal of Educational Psychology, 2*, 520–531. doi:10.1037/edu0000149

Xenidou-Dervou, I., Van Luit, J. E. H., Kroesbergen, E. H., Friso-van den Bos, I., Jonkman, L. M., & van der Schoot, M., … van Lieshout, C. D. M. (2018). Cognitive predictors of children’s development in mathematics achievement: A latent growth modeling approach. *Developmental Science,* 21, 1–14. doi:10.1111/desc.12671

Yalçin, S., & Tavşancil, E. (2014). The comparison of Turkish students' PISA achievement levels by year via data envelopment analysis. *Educational Sciences: Theory & Practice, 14*(3), 961–968. doi:10.12738/estp.2014.3.1748

Yamamoto, K., & Lennon, M. L. (2017). Understanding and detecting data fabrication in large-scale assessments. *Quality Assurance in Education 26* (2), 196–212. Doi:10.1108/QAE-07-2017-0038

You, Y. (2018). The seeming 'round trip' of learner-centered education: A 'best practice' derived from China's new curriculum reform? *Journal of Comparative Education 54*, 97–115. doi:10.1080/03050068.2018.1541662.

Želvys, R. (2017). The PISA phenomenon: The many faces of international student assessment. *Acta Paedagogica Vilnensia, 37*. doi:10.15388/ActPaed.2017.37.10462

Zhang, D., Ding, Y., Stegall, J., & Mo, L. (2012). The effect of visual-chunking-representation accommodation on geometry testing for students with math disabilities. *Learning Disabilities Research & Practice*, *27*(4), 167–177. doi:10.1111/j.1540-5826.2012.00364.x

Zhang, X., Räsänen, P., Koponen, T., Aunola, K., Lerkkanen, M., & Nurmi, J. (2017). Knowing, applying and reasoning about arithmetic: Roles of domain general and numerical skills in multiple domains of arithmetic learning. *Developmental Psychology, 53*(12), 2304–2318. doi:10.1037/dev0000432

Appendices

Background: In this study, I used archival data which comprised Vincentian students' scores in the CSEC mathematic examination, and statistics (means) on CXC mathematics examination. The data were used at two points in the study. Initially, data was required at the proposal stage to establish the problem under study. Once approved, access to data was requested to answer the research question. Permission for the use of the data was sought and granted at two different points in time, and from two entities. One entity was the Ministry of Education, National Reconciliation and Information, Saint Vincent and the Grenadines, which is the owner of the data. The other entity was the CXC, which is the custodian of the data.

Contents: The following appendices, A–G, indicate the various request for and approval of the use of students' examination data.

Appendix A: Request for permission to access and use the CXC mathematics Examination data – Letter to registrar [preliminary proposal-level data analysis]

Appendix B: Approval to access and use the CXC mathematics Examination data – Letter from registrar.[preliminary proposal-level data analysis]

Appendix C: Request for permission to access and use Vincentian students' scores in the CSEC mathematics Examination – Letter to registrar

Appendix D: Approval to access and use Vincentian students' scores in the CSEC mathematics Examination – Letter from registrar

Appendix E: Request for permission to access and use Vincentian students' scores in the CSEC mathematics Examination – Letter to Permanent Secretary

Appendix A: Request for Permission To Access and Use the CXC Mathematics Examination Data – Letter to Registrar

141 Flamboyant Avenue
Warners Park
Christ Church

The Registrar
Caribbean Examinations Council
Prince Road, Pine Plantation Road
St. Michael

16th January, 2018

Dear Sir

I am currently pursuing a PhD in Assessment, Evaluation and Accountability with Walden University. I have completed all my course work requirements and I am now at the dissertation stage. From my experience as a teacher and my position as assessment officer in the organization, I am aware that there is a problem with the mathematics achievement of students in the Caribbean. Every year, less than 45 percent of the students achieve acceptable grades on the Caribbean Secondary Education Certificate (CSEC) mathematics examination. The mathematics result is usually a topic of discussion among ministers of education throughout the region. However, despite efforts to improve, the poor performance persists.

Mathematics is important in every sphere of life, including entry level job requirement, access to higher education and better career options such as medicine, engineering and statistics. Given the prestige that is associated with mathematics, this area needs urgent attention, including empirical research to explore the sources of the problem and provide recommendations for improvement. With my passion for mathematics, and in keeping with the organization's drive to encourage staff engagement in research, I hereby seek your permission for the use of the Council's data to investigate the problem as part of the requirement for the fulfilment of my dissertation. The required data will include subject means for five years, and question means for the May/June 2017 CSEC Mathematics examination in order to establish that a problem exists and is worthy of investigation. In addition, I will like to review past CSEC mathematics Paper 02 questions to guide my construction of a mathematics test that will be administered to a sample of students.

I assure the Council that anonymity of candidates' identity will be observed and all sensitive information will be treated with the strictest confidence. I believe that the proposed research will not only benefit me on a personal and professional basis, but the Council and the wider Caribbean will also benefit. The findings will provide invaluable insights into the problem and recommendations for possible remediation.

Thanking you in advance for your continued support of higher learning and professional development.

Yours Respectfully

BRCato

Brendalee Cato

Approval Granted

[illegible]

07.3.2018

Appendix B: Approval To Access and Use the CXC Mathematics Examination Data – Letter From Registrar

CARIBBEAN EXAMINATIONS COUNCIL

HEADQUARTERS
Prince Road, Pine Plantation Road,
St. Michael BB11091, Barbados
t: +1 (246) 227-1700 f: +1 (246) 429-5421
e: cxcezo@cxc.org
w: www.cxc.org | www.cxc-store.com

In reply please quote our ref:

27 March 2018

Ms Brendalee Cato
#141 Flamboyant Avenue
CHRIST CHURCH

Dear Ms Cato

Re: Doctoral Studies in Assessment, Evaluation and Accountability – Walden University

I am in receipt of your letter dated 16 January 2018 in respect of the captioned subject.

This correspondence serves to confirm that permission is granted for you to use data or information, the property of the Caribbean Examinations Council, to research the problem related to the mathematics achievement of students in the Caribbean.

Permission is being given with the understanding that all sensitive or identifiable data and information belonging to the Caribbean Examinations Council or which comes into the possession of the Council through its work with Ministries of Education across the CXC Participating Territories, candidates and any other agents, will be treated responsibly and in the strictest of confidence.

I wish you success in your studies.

Yours sincerely

Glenroy Cumberbatch
REGISTRAR and CEO

GC/amg

CPEA™ CCSLC® CVQ* CSEC® CAPE® CXC®-AD

*CVQ is the trademark of the Caribbean Association of National Training Authorities

Appendix C: Request for Permission To Access and Use Vincentian Students' Scores In the CSEC Mathematics Examination – Letter to Registrar

141 Flamboyant Avenue
Warners Park
Christ Church

The Registrar
Caribbean Examinations Council
Prince Road, Pine Plantation Road
St. Michael

13 February, 2020

Dear Sir
I am currently pursuing a PhD in Assessment, Evaluation and Accountability at Walden University. I have completed all my course work requirements and I am now at the dissertation stage. From my experience as a teacher and my position as assessment officer in the organization, I am aware that there is a problem with the mathematics achievement of students in the Caribbean. Every year, less than 45 percent of the students achieve acceptable grades on the Caribbean Secondary Education Certificate (CSEC) mathematics examination. The mathematics result is usually a topic of discussion among ministers of education throughout the region. However, despite efforts to improve, the poor performance persists.

Mathematics is important in every sphere of life, including entry level job requirement, access to higher education and better career options such as medicine, engineering and statistics. Given the high esteem and prestige that is associated with mathematics, this area needs urgent attention, including empirical research to explore the problem and provide recommendations for improvement. With my passion for mathematics, and in keeping with the organization's drive to encourage staff engagement in research, I hereby seek your permission for the use of the Council's data to investigate the problem and fulfil my dissertation requirements. The required data will include questions scores by content and cognitive strand for Vincentian students in the 2017 May/June CSEC examination. Students names or registration numbers are not required.

I assure the Council that anonymity of candidates' identity and data will be observed and all information will be treated with the strictest confidence. I believe that the proposed research will not only benefit me on a personal and professional basis, but the findings will be beneficial to the Council and the wider Caribbean. I hope to provide recommendations that should help education policy makers devise strategies for remedial action, thereby influencing positive social chance in the Caribbean.

In order to achieve ethics and compliance approval for the collection of data, Walden University requires you to endorse the attached data use agreement form.

Thanking you in advance for your support of higher learning and professional development.

Yours Respectfully

BRCato

Brendalee Cato

Appendix D: Approval To Access and Use Vincentian Students' Scores in the CSEC Mathematics Examination – Letter From Registrar

CARIBBEAN EXAMINATIONS COUNCIL

HEADQUARTERS
Prince Road, Pine Plantation Road
St. Michael BB11091, Barbados
t: +1 (246) 227-1700 f: +1 (246) 429-5421
e: cxcezo@cxc.org
w: www.cxc.org | www.cxc-store.com

In reply please quote our ref: 0485-0284P

20 March 2020

Ms Brendalee Cato
141 Flamboyant Avenue
Warners Park
CHRIST CHURCH

Dear Ms Cato

Permission to Use Candidate Data

Reference is made to your correspondence dated 13 February 2020 requesting permission to use Mathematics data for Vincentian students from the 2017 May-June examinations, which would enable you to complete research towards the fulfillment of your dissertation.

CXC notes the "no objection letter" from the Ministry of Education, National Reconciliation and Information St Vincent and the Grenadines for access to their country specific data.

Approval is hereby given, as per letter dated 5 March 2020 from the Permanent Secretary Mr Mycele Burke, Ministry of Education National Reconciliation and Information St Vincent and the Grenadines (copy attached).

Best wishes for successful completion of your PhD in Assessment, Evaluation and Accountability.

Yours sincerely

Wayne Wesley, PhD, CMgr
REGISTRAR and CEO

/amg

att

Appendix E: Request for Permission To Access and Use Vincentian Students' Scores in the CSEC Mathematics Examination – Letter to Permanent Secretary

141 Flamboyant Avenue
Warners Park
Christ Church

The Permanent Secretary (ag.)
Ministry of Education National Reconciliation and Information
Halifax Street, Kingstown
Saint Vincent and the Grenadines

14 February, 2020

Dear Sir
I am currently pursuing a PhD in Assessment, Evaluation and Accountability at Walden University. For my dissertation, I am investigating the poor mathematics performance of Vincentian students in the Caribbean Secondary Examinations Certificate (CSEC) mathematics examination. Specifically, I would be analyzing students' performance in the 2017 May/June CSEC examination by content and cognitive strand (domain). While students' performance by question is required, the identity of the students is not required. The data to be used in the study will be archival data, currently housed at the Caribbean Examinations Council. While the Council supports the study, your approval of the use of the data is required.

I assure the ministry of education that anonymity of the students' identity and data will be observed and all information will be treated with the strictest confidence. I hope through this research to provide recommendations to your ministry that should help in policy decisions regarding strategies for improving mathematics performance in Saint Vincent and the Grenadines and by extension, the Caribbean.

Thanking you in advance for your support of this worthwhile venture.

Yours Respectfully

BR-Cato

Brendalee Cato

Appendix F: Approval to Access and Use Vincentian Students' Scores in the CSEC Mathematics Examination – Letter From Permanent Secretary

Ref No:
In replying the date and number above of this letter should be quoted

Tel. No. 1 (784) 457-1104 or 457-2576
Fax No: 1 (784) 457-1114

MINISTRY OF EDUCATION, NATIONAL RECONCILIATION AND INFORMATION

Halifax Street, Kingstown
St. Vincent and the Grenadines

5th March, 2020

Ms. Brendalee Cato
Manager
Examinations Development and Production Division
Caribbean Examinations Council
Prince Road, Pine Plantation Road,
St. Michael BB11091,
BARBADOS

Dear Ms. Cato,

The Ministry of Education, National Reconciliation and Information will like to congratulate you on pursing the PhD in Assessment, Evaluation and Accountability at the Walden University.

The Ministry of Education, National Reconciliation and Information of St. Vincent and the Grenadines has no objection to you analysing Vincentian students' question level scores in the May/June 2017 Caribbean Secondary Examination Certificate (CSEC) mathematics examination as part of your doctoral studies.

Please be informed that your request to investigate the poor Mathematics performance of Vincentian students in the Caribbean Secondary Education Certificate (CSEC) Examination was received on February 14, 2020. The Ministry of Education has no objection to you conducting this research.

The Ministry of Education, National Reconciliation and Information sincerely wishes you success on your studies.

Yours Sincerely,

MINISTRY OF EDUCATION
NATIONAL RECONCILIATION
AND INFORMATION
ST. VINCENT & THE GRENADINES

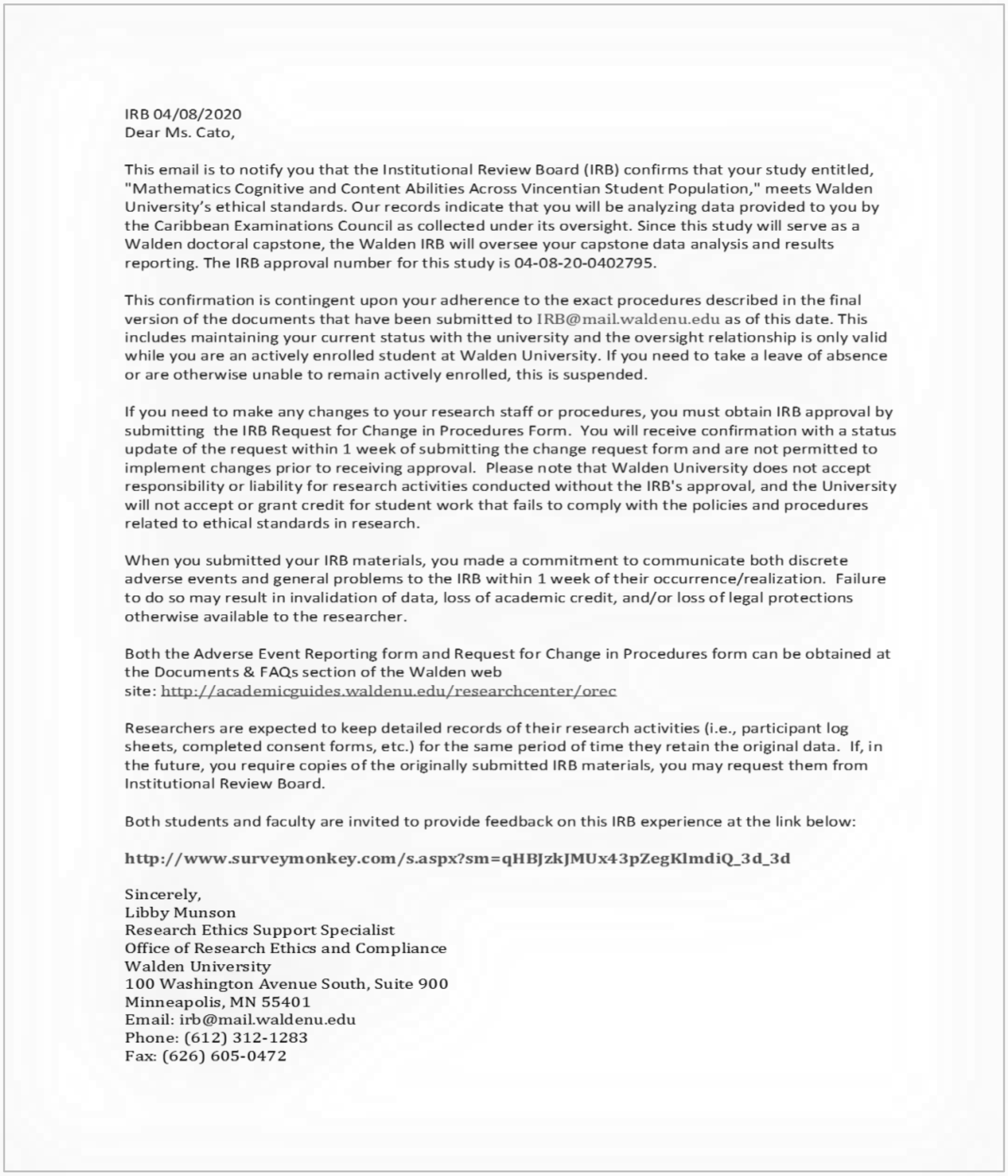
....
Mycsie Burke (Mrs.)
Permanent Secretary
Ministry of Education, National Reconciliation and Information

---

Email:

Administration:	office.education@mail.gov.vc	UNESCO:	unescosvg@gov.vc
Minister:	minister.emen@gmail.com	EXAMS:	moeexamsunit@gmail.com
Permanent Secretary:	pseducation@outlook.com	Curriculum Unit:	svgcurriculumunit@gmail.com
Chief Education Officer:	ceo.education@mail.gov.vc	Adult and Continuing Education Unit:	adulteducation2@yahoo.com

Appendix G: Approval From Walden's IRB To Collect Data



IRB 04/08/2020
Dear Ms. Cato,

This email is to notify you that the Institutional Review Board (IRB) confirms that your study entitled, "Mathematics Cognitive and Content Abilities Across Vincentian Student Population," meets Walden University's ethical standards. Our records indicate that you will be analyzing data provided to you by the Caribbean Examinations Council as collected under its oversight. Since this study will serve as a Walden doctoral capstone, the Walden IRB will oversee your capstone data analysis and results reporting. The IRB approval number for this study is 04-08-20-0402795.

This confirmation is contingent upon your adherence to the exact procedures described in the final version of the documents that have been submitted to IRB@mail.waldenu.edu as of this date. This includes maintaining your current status with the university and the oversight relationship is only valid while you are an actively enrolled student at Walden University. If you need to take a leave of absence or are otherwise unable to remain actively enrolled, this is suspended.

If you need to make any changes to your research staff or procedures, you must obtain IRB approval by submitting the IRB Request for Change in Procedures Form. You will receive confirmation with a status update of the request within 1 week of submitting the change request form and are not permitted to implement changes prior to receiving approval. Please note that Walden University does not accept responsibility or liability for research activities conducted without the IRB's approval, and the University will not accept or grant credit for student work that fails to comply with the policies and procedures related to ethical standards in research.

When you submitted your IRB materials, you made a commitment to communicate both discrete adverse events and general problems to the IRB within 1 week of their occurrence/realization. Failure to do so may result in invalidation of data, loss of academic credit, and/or loss of legal protections otherwise available to the researcher.

Both the Adverse Event Reporting form and Request for Change in Procedures form can be obtained at the Documents & FAQs section of the Walden web site: http://academicguides.waldenu.edu/researchcenter/orec

Researchers are expected to keep detailed records of their research activities (i.e., participant log sheets, completed consent forms, etc.) for the same period of time they retain the original data. If, in the future, you require copies of the originally submitted IRB materials, you may request them from Institutional Review Board.

Both students and faculty are invited to provide feedback on this IRB experience at the link below:

http://www.surveymonkey.com/s.aspx?sm=qHBJzkJMUx43pZegKlmdiQ_3d_3d

Sincerely,
Libby Munson
Research Ethics Support Specialist
Office of Research Ethics and Compliance
Walden University
100 Washington Avenue South, Suite 900
Minneapolis, MN 55401
Email: irb@mail.waldenu.edu
Phone: (612) 312-1283
Fax: (626) 605-0472

Appendix H: Request for Vincentian Students' Scores in the CSEC Mathematics Examination – Email to Data Manager

From: Brendalee Cato
Sent: Friday, 10 April 2020 13:34
To: Andre Blair <ABlair@cxc.org>
Subject: Request for data

Good afternoon Andre. I hope you and your family are keeping safe amidst COVID-19.

On Wednesday 8th April, 2020, I received approval from the university Institutional Review Board (IRB) to proceed with data collection. For my research I am using archival data comprising the scores (by content domain and cognitive domain) of students from Saint Vincent and the Grenadines in the 2017 May/June CSEC mathematics examination.

I have attached the following documents for your guidance

- Letter of approval for data usage from the Caribbean Examinations Council
- Letter of approval for use of data from Saint Vincent and the Grenadines
- Table indicating how the data should be organized

I will greatly appreciate if this request could be honoured within the coming week.

Thanks in Advance

Brendalee

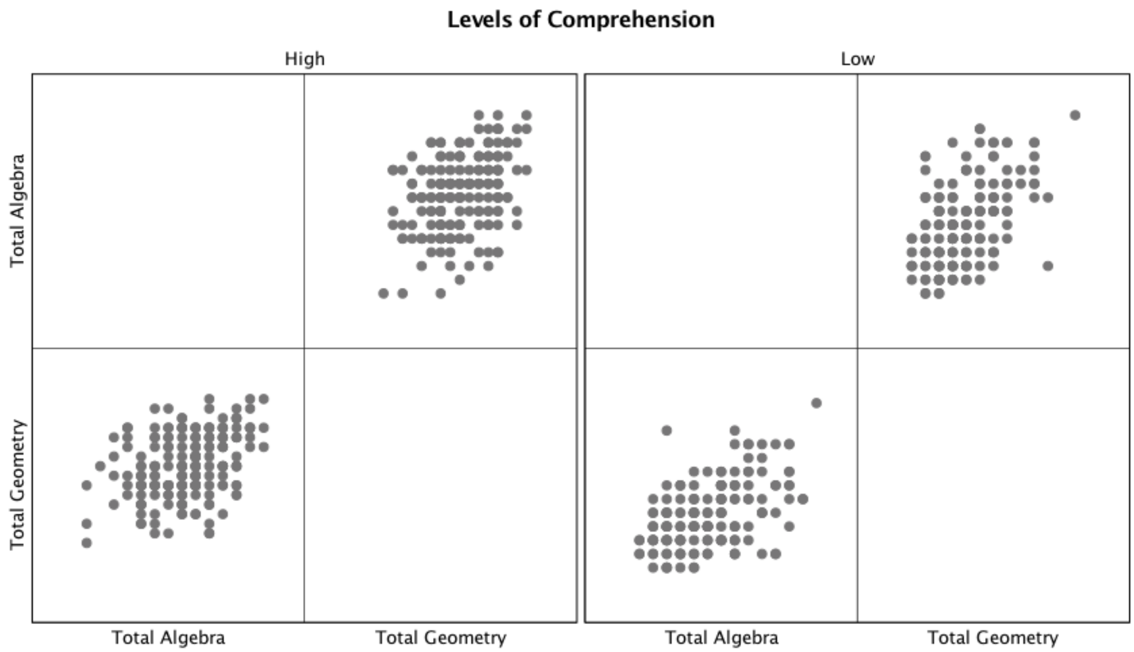
Brendalee Cato
Manager
Examinations Development and Production Division

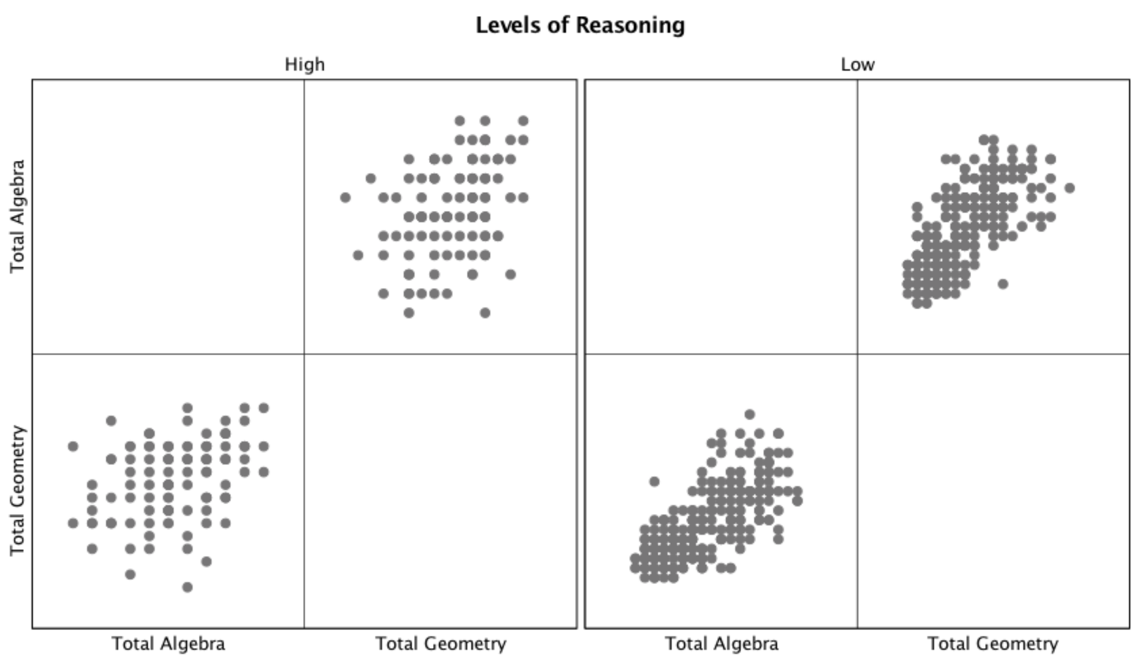
Caribbean Examinations Council
Prince Road, Pine Plantation Road,
St. Michael BB11091, Barbados

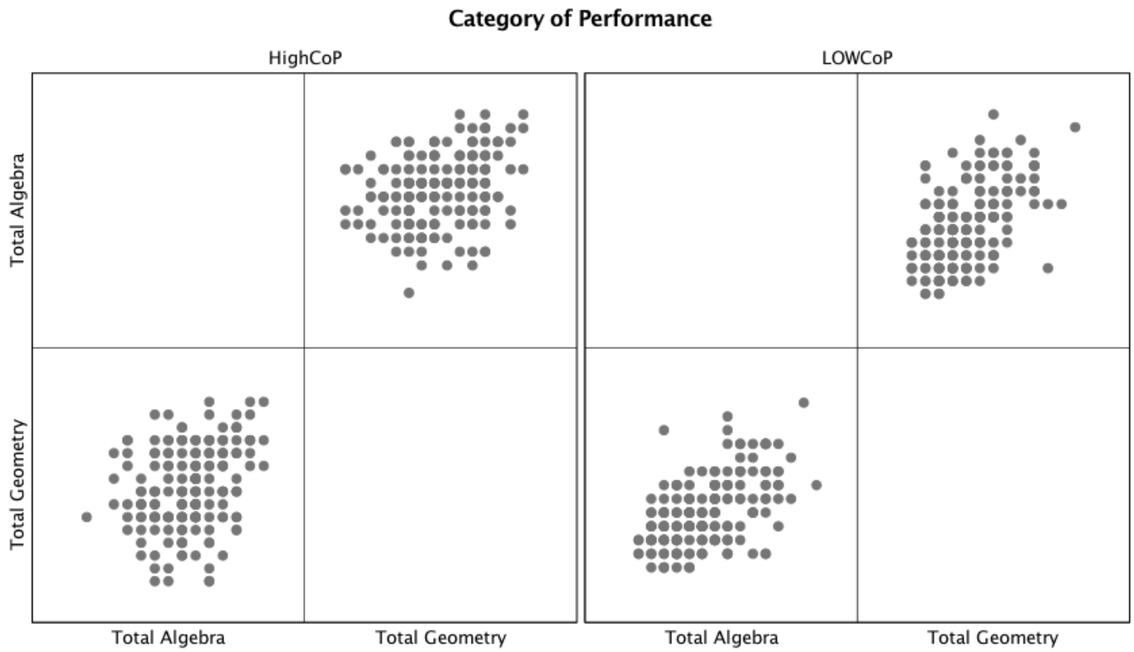
t: +1 (246) 227-1843 f: +1 (246) 429-5421
e: cxcezo@cxc.org w: www.cxc.org | www.cxc-store.com

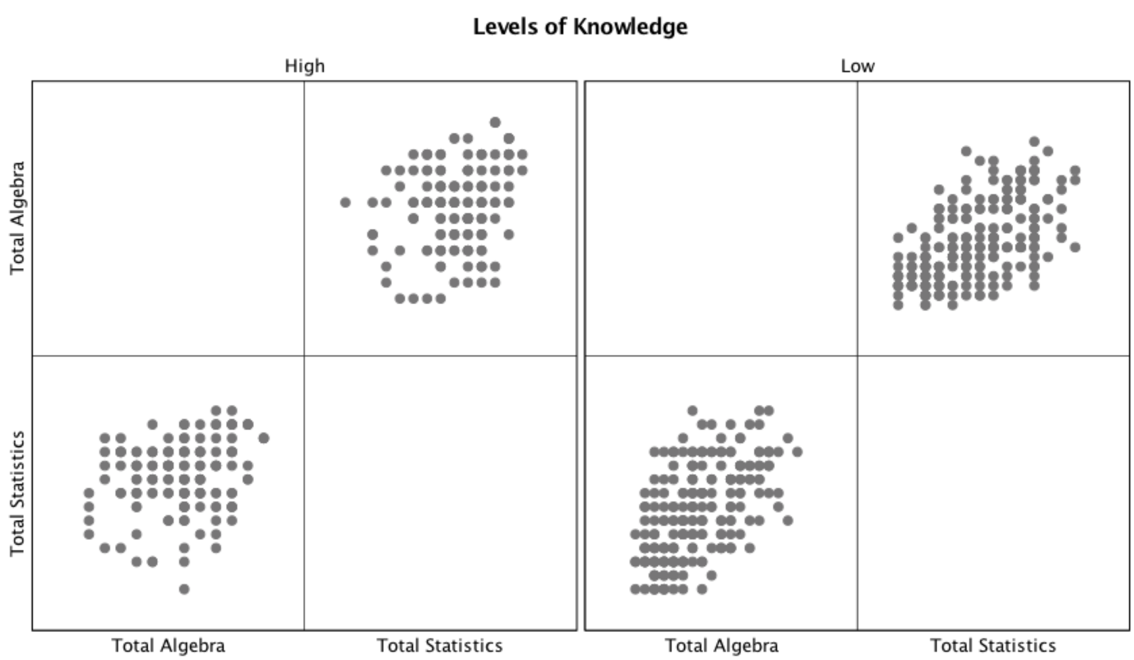
3 Attachments

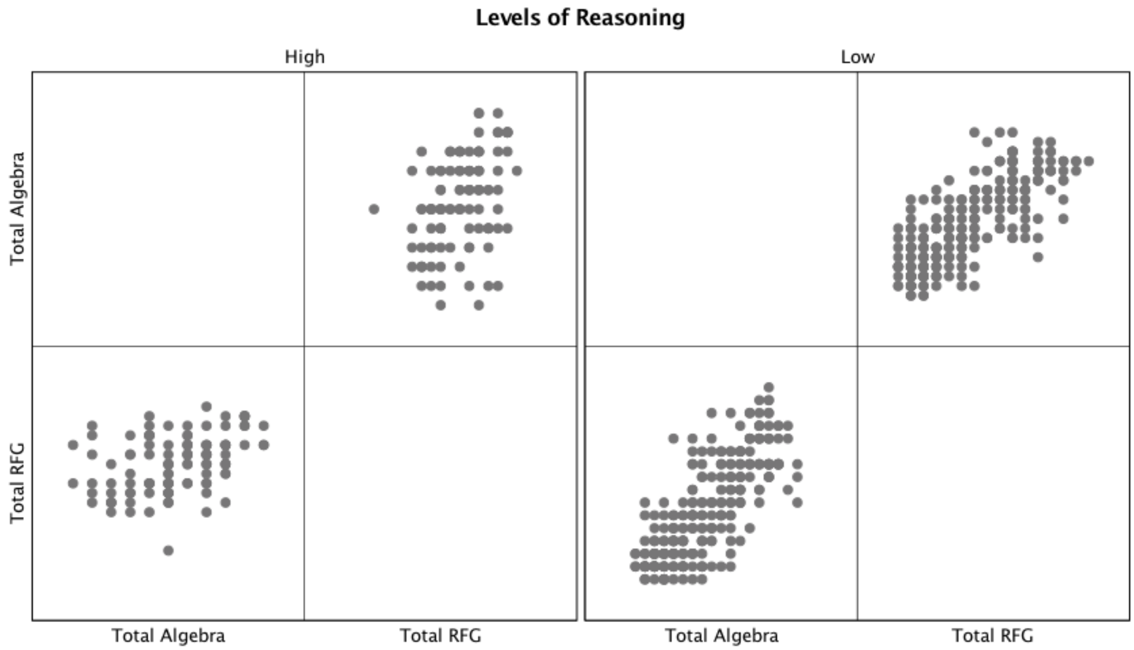
Appendix I: Scatterplots – Pairs of Dependent Variables and Levels of Independent Variables

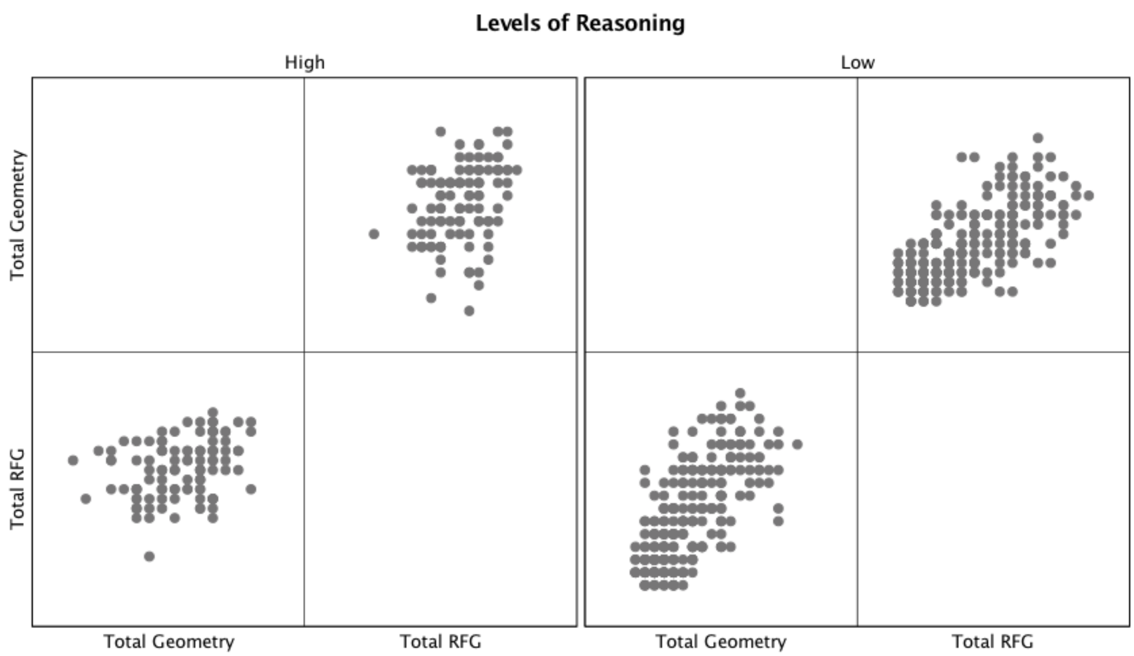




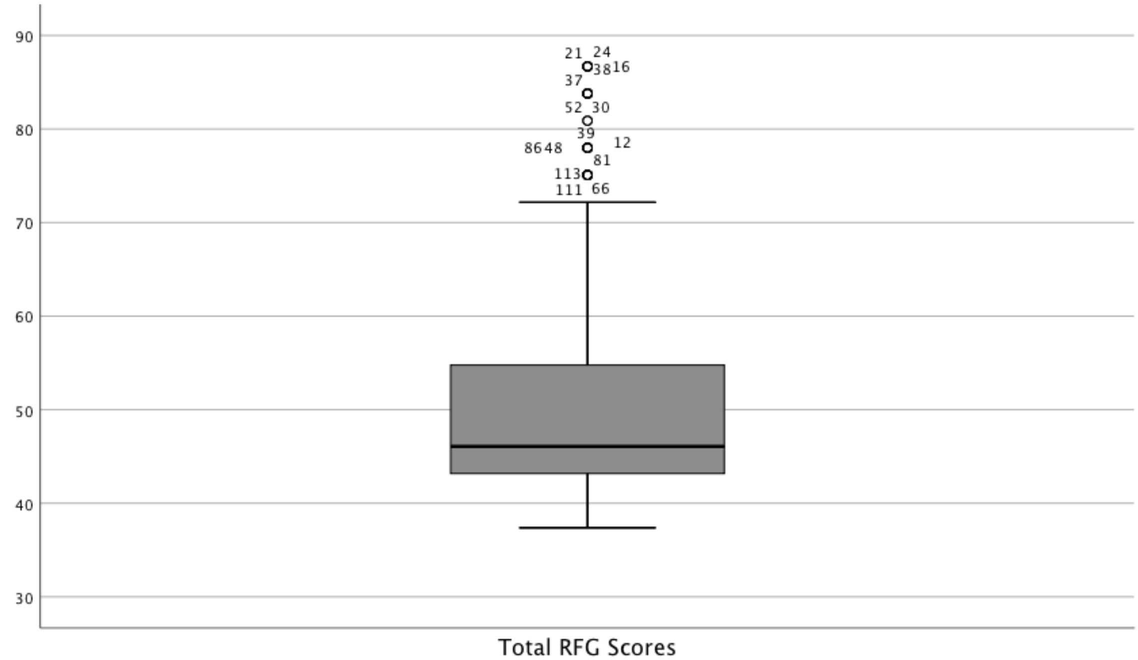

Category of Performance
HighCoP
LOWCoP
Total Algebra
Total Geometry
Total Algebra
Total Geometry
Total Algebra
Total Geometry

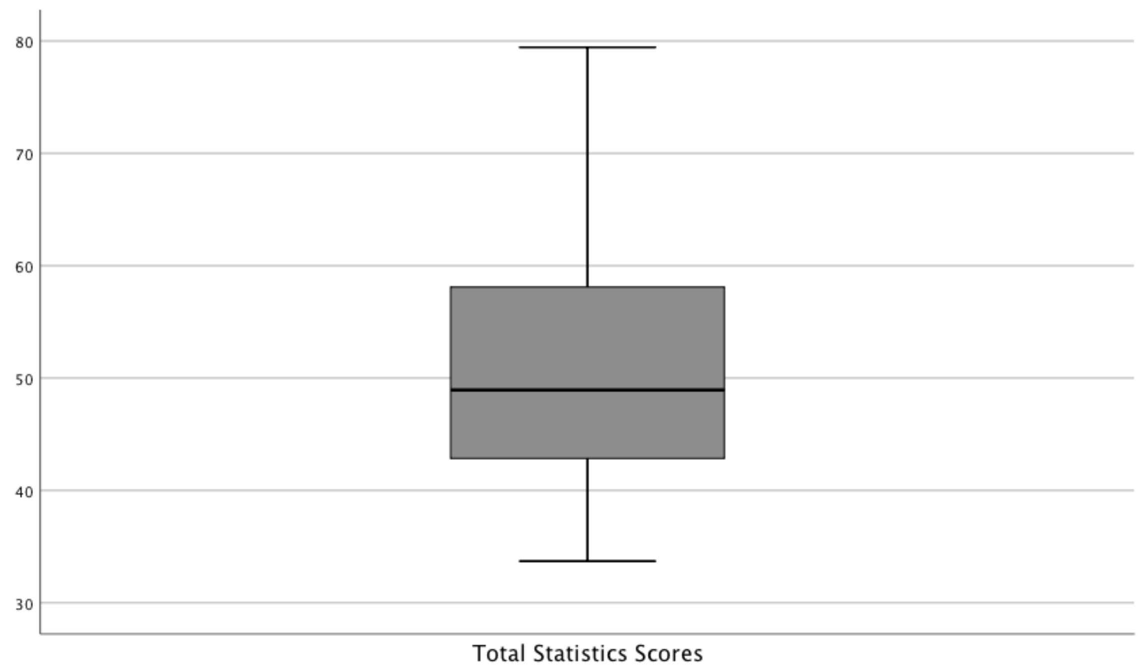

Levels of Knowledge
High
Low
Total Algebra
Total Statistics
Total Algebra
Total Statistics
Total Algebra
Total Statistics


Levels of Reasoning
High
Low
Total Algebra
Total RFG
Total Algebra
Total RFG
Total Algebra
Total RFG


Levels of Reasoning
High
Low
Total Geometry
Total RFG
Total Geometry
Total RFG
Total Geometry
Total RFG

Appendix J: Sample Boxplots





Appendix K: Descriptive Statistic

Descriptive Statistics

	Levels of Knowledge	Levels of Comprehension	Levels of Reasoning	Category of Performance	Mean	Std. Deviation	N
	Low	Low	Low	LOWCoP	5.90	2.847	168
				Total	5.90	2.847	168
			High	LOWCoP	12.00	.	1
				Total	12.00	.	1
			Total	LOWCoP	5.93	2.877	169
				Total	5.93	2.877	169
		High	Low	LOWCoP	10.46	2.259	13
				HighCoP	13.19	2.354	27
				Total	12.30	2.633	40
			High	HighCoP	14.08	2.109	12
				Total	14.08	2.109	12
			Total	LOWCoP	10.46	2.259	13
				HighCoP	13.46	2.292	39
				Total	12.71	2.615	52
		Total	Low	LOWCoP	6.23	3.042	181
				HighCoP	13.19	2.354	27
				Total	7.13	3.773	208
			High	LOWCoP	12.00	.	1
				HighCoP	14.08	2.109	12
				Total	13.92	2.100	13
			Total	LOWCoP	6.26	3.064	182
				HighCoP	13.46	2.292	39
				Total	7.53	4.026	221
	High	Low	Low	LOWCoP	10.00	1.732	3
				Total	10.00	1.732	3
			High	HighCoP	12.50	.707	2
				Total	12.50	.707	2
			Total	LOWCoP	10.00	1.732	3
				HighCoP	12.50	.707	2

			Total	11.00	1.871	5
	High	Low	HighCoP	13.68	2.260	53
			Total	13.68	2.260	53
		High	HighCoP	15.65	2.392	91
			Total	15.65	2.392	91
		Total	HighCoP	14.92	2.523	144
			Total	14.92	2.523	144
	Total	Low	LOWCoP	10.00	1.732	3
			HighCoP	13.68	2.260	53
			Total	13.48	2.374	56
		High	HighCoP	15.58	2.411	93
			Total	15.58	2.411	93
		Total	LOWCoP	10.00	1.732	3
			HighCoP	14.89	2.522	146
			Total	14.79	2.597	149
Total	Low	Low	LOWCoP	5.97	2.879	171
			Total	5.97	2.879	171
		High	LOWCoP	12.00	.	1
			HighCoP	12.50	.707	2
			Total	12.33	.577	3
		Total	LOWCoP	6.01	2.907	172
			HighCoP	12.50	.707	2
			Total	6.08	2.973	174
	High	Low	LOWCoP	10.46	2.259	13
			HighCoP	13.51	2.289	80
			Total	13.09	2.509	93
		High	HighCoP	15.47	2.404	103
			Total	15.47	2.404	103
		Total	LOWCoP	10.46	2.259	13
			HighCoP	14.61	2.541	183
			Total	14.34	2.723	196
	Total	Low	LOWCoP	6.29	3.060	184
			HighCoP	13.51	2.289	80
			Total	8.48	4.377	264
		High	LOWCoP	12.00	.	1
			HighCoP	15.41	2.417	105
			Total	15.38	2.428	106
		Total	LOWCoP	6.32	3.081	185

				HighCoP	14.59	2.538	185
				Total	10.45	5.009	370
Total Geometry	Low	Low	Low	LOWCoP	4.65	2.275	168
				Total	4.65	2.275	168
			High	LOWCoP	10.00	.	1
				Total	10.00	.	1
			Total	LOWCoP	4.68	2.305	169
				Total	4.68	2.305	169
		High	Low	LOWCoP	8.00	2.198	13
				HighCoP	9.59	2.859	27
				Total	9.08	2.740	40
			High	HighCoP	11.50	2.153	12
				Total	11.50	2.153	12
			Total	LOWCoP	8.00	2.198	13
				HighCoP	10.18	2.780	39
				Total	9.63	2.794	52
		Total	Low	LOWCoP	4.89	2.424	181
				HighCoP	9.59	2.859	27
				Total	5.50	2.941	208
			High	LOWCoP	10.00	.	1
				HighCoP	11.50	2.153	12
				Total	11.38	2.103	13
			Total	LOWCoP	4.92	2.447	182
				HighCoP	10.18	2.780	39
				Total	5.85	3.210	221
	High	Low	Low	LOWCoP	8.67	1.528	3
				Total	8.67	1.528	3
			High	HighCoP	8.00	2.828	2
				Total	8.00	2.828	2
			Total	LOWCoP	8.67	1.528	3
				HighCoP	8.00	2.828	2
				Total	8.40	1.817	5
		High	Low	HighCoP	11.58	2.583	53
				Total	11.58	2.583	53
			High	HighCoP	13.84	2.918	91
				Total	13.84	2.918	91
			Total	HighCoP	13.01	2.995	144
				Total	13.01	2.995	144

		Total	Low	LOWCoP	8.67	1.528	3
				HighCoP	11.58	2.583	53
				Total	11.43	2.614	56
			High	HighCoP	13.71	3.024	93
				Total	13.71	3.024	93
			Total	LOWCoP	8.67	1.528	3
				HighCoP	12.94	3.040	146
				Total	12.85	3.074	149
	Total	Low	Low	LOWCoP	4.72	2.322	171
				Total	4.72	2.322	171
			High	LOWCoP	10.00	.	1
				HighCoP	8.00	2.828	2
				Total	8.67	2.309	3
			Total	LOWCoP	4.75	2.350	172
				HighCoP	8.00	2.828	2
				Total	4.79	2.372	174
		High	Low	LOWCoP	8.00	2.198	13
				HighCoP	10.91	2.825	80
				Total	10.51	2.918	93
			High	HighCoP	13.56	2.929	103
				Total	13.56	2.929	103
			Total	LOWCoP	8.00	2.198	13
				HighCoP	12.40	3.164	183
				Total	12.11	3.294	196
		Total	Low	LOWCoP	4.95	2.457	184
				HighCoP	10.91	2.825	80
				Total	6.76	3.759	264
			High	LOWCoP	10.00	.	1
				HighCoP	13.46	3.013	105
				Total	13.42	3.017	106
			Total	LOWCoP	4.98	2.478	185
				HighCoP	12.36	3.187	185
				Total	8.67	4.666	370
Total Measurement	Low	Low	Low	LOWCoP	3.63	2.058	168
				Total	3.63	2.058	168
			High	LOWCoP	12.00	.	1
				Total	12.00	.	1
			Total	LOWCoP	3.68	2.150	169

			Total	3.68	2.150	169
	High	Low	LOWCoP	9.08	3.818	13
			HighCoP	9.81	3.476	27
			Total	9.58	3.558	40
		High	HighCoP	9.17	2.791	12
			Total	9.17	2.791	12
		Total	LOWCoP	9.08	3.818	13
			HighCoP	9.62	3.258	39
			Total	9.48	3.375	52
	Total	Low	LOWCoP	4.02	2.625	181
			HighCoP	9.81	3.476	27
			Total	4.77	3.364	208
		High	LOWCoP	12.00	.	1
			HighCoP	9.17	2.791	12
			Total	9.38	2.785	13
		Total	LOWCoP	4.07	2.683	182
			HighCoP	9.62	3.258	39
			Total	5.05	3.500	221
High	Low	Low	LOWCoP	11.33	2.082	3
			Total	11.33	2.082	3
		High	HighCoP	8.50	3.536	2
			Total	8.50	3.536	2
		Total	LOWCoP	11.33	2.082	3
			HighCoP	8.50	3.536	2
			Total	10.20	2.775	5
	High	Low	HighCoP	10.13	2.781	53
			Total	10.13	2.781	53
		High	HighCoP	13.99	3.692	91
			Total	13.99	3.692	91
		Total	HighCoP	12.57	3.857	144
			Total	12.57	3.857	144
	Total	Low	LOWCoP	11.33	2.082	3
			HighCoP	10.13	2.781	53
			Total	10.20	2.746	56
		High	HighCoP	13.87	3.757	93
			Total	13.87	3.757	93
		Total	LOWCoP	11.33	2.082	3
			HighCoP	12.51	3.871	146

				Total	12.49	3.843	149
	Total	Low	Low	LOWCoP	3.77	2.289	171
				Total	3.77	2.289	171
			High	LOWCoP	12.00	.	1
				HighCoP	8.50	3.536	2
				Total	9.67	3.215	3
			Total	LOWCoP	3.81	2.367	172
				HighCoP	8.50	3.536	2
				Total	3.87	2.421	174
		High	Low	LOWCoP	9.08	3.818	13
				HighCoP	10.02	3.015	80
				Total	9.89	3.133	93
			High	HighCoP	13.43	3.910	103
				Total	13.43	3.910	103
			Total	LOWCoP	9.08	3.818	13
				HighCoP	11.94	3.921	183
				Total	11.75	3.970	196
		Total	Low	LOWCoP	4.14	2.772	184
				HighCoP	10.02	3.015	80
				Total	5.92	3.926	264
			High	LOWCoP	12.00	.	1
				HighCoP	13.33	3.946	105
				Total	13.32	3.929	106
			Total	LOWCoP	4.18	2.824	185
				HighCoP	11.90	3.925	185
				Total	8.04	5.157	370
Total Statistics	Low	Low	Low	LOWCoP	4.66	3.028	168
				Total	4.66	3.028	168
			High	LOWCoP	4.00	.	1
				Total	4.00	.	1
			Total	LOWCoP	4.66	3.020	169
				Total	4.66	3.020	169
		High	Low	LOWCoP	7.62	2.987	13
				HighCoP	8.63	2.169	27
				Total	8.30	2.472	40
			High	HighCoP	9.75	1.960	12
				Total	9.75	1.960	12
			Total	LOWCoP	7.62	2.987	13

			HighCoP	8.97	2.146	39
			Total	8.63	2.426	52
	Total	Low	LOWCoP	4.87	3.113	181
			HighCoP	8.63	2.169	27
			Total	5.36	3.259	208
		High	LOWCoP	4.00	.	1
			HighCoP	9.75	1.960	12
			Total	9.31	2.463	13
		Total	LOWCoP	4.87	3.105	182
			HighCoP	8.97	2.146	39
			Total	5.59	3.345	221
High	Low	Low	LOWCoP	5.67	1.528	3
			Total	5.67	1.528	3
		High	HighCoP	11.50	.707	2
			Total	11.50	.707	2
		Total	LOWCoP	5.67	1.528	3
			HighCoP	11.50	.707	2
			Total	8.00	3.391	5
	High	Low	HighCoP	9.09	2.115	53
			Total	9.09	2.115	53
		High	HighCoP	10.97	2.496	91
			Total	10.97	2.496	91
		Total	HighCoP	10.28	2.524	144
			Total	10.28	2.524	144
	Total	Low	LOWCoP	5.67	1.528	3
			HighCoP	9.09	2.115	53
			Total	8.91	2.218	56
		High	HighCoP	10.98	2.471	93
			Total	10.98	2.471	93
		Total	LOWCoP	5.67	1.528	3
			HighCoP	10.29	2.511	146
			Total	10.20	2.576	149
Total	Low	Low	LOWCoP	4.68	3.009	171
			Total	4.68	3.009	171
		High	LOWCoP	4.00	.	1
			HighCoP	11.50	.707	2
			Total	9.00	4.359	3
		Total	LOWCoP	4.67	3.001	172

				HighCoP	11.50	.707	2
				Total	4.75	3.072	174
		High	Low	LOWCoP	7.62	2.987	13
				HighCoP	8.94	2.131	80
				Total	8.75	2.297	93
			High	HighCoP	10.83	2.463	103
				Total	10.83	2.463	103
			Total	LOWCoP	7.62	2.987	13
				HighCoP	10.00	2.501	183
				Total	9.84	2.596	196
		Total	Low	LOWCoP	4.89	3.093	184
				HighCoP	8.94	2.131	80
				Total	6.11	3.391	264
			High	LOWCoP	4.00	.	1
				HighCoP	10.84	2.442	105
				Total	10.77	2.520	106
			Total	LOWCoP	4.88	3.085	185
				HighCoP	10.02	2.492	185
				Total	7.45	3.802	370
Total RFG	Low	Low	Low	LOWCoP	3.21	2.380	168
				Total	3.21	2.380	168
			High	LOWCoP	8.00	.	1
				Total	8.00	.	1
			Total	LOWCoP	3.24	2.401	169
				Total	3.24	2.401	169
		High	Low	LOWCoP	7.08	1.977	13
				HighCoP	9.00	2.075	27
				Total	8.38	2.215	40
			High	HighCoP	9.92	2.746	12
				Total	9.92	2.746	12
			Total	LOWCoP	7.08	1.977	13
				HighCoP	9.28	2.305	39
				Total	8.73	2.410	52
		Total	Low	LOWCoP	3.49	2.553	181
				HighCoP	9.00	2.075	27
				Total	4.20	3.108	208
			High	LOWCoP	8.00	.	1
				HighCoP	9.92	2.746	12

			Total	9.77	2.682	13
		Total	LOWCoP	3.51	2.568	182
			HighCoP	9.28	2.305	39
			Total	4.53	3.347	221
High	Low	Low	LOWCoP	7.67	3.055	3
			Total	7.67	3.055	3
		High	HighCoP	9.50	.707	2
			Total	9.50	.707	2
		Total	LOWCoP	7.67	3.055	3
			HighCoP	9.50	.707	2
			Total	8.40	2.408	5
	High	Low	HighCoP	10.40	2.106	53
			Total	10.40	2.106	53
		High	HighCoP	12.48	2.884	91
			Total	12.48	2.884	91
		Total	HighCoP	11.72	2.805	144
			Total	11.72	2.805	144
	Total	Low	LOWCoP	7.67	3.055	3
			HighCoP	10.40	2.106	53
			Total	10.25	2.218	56
		High	HighCoP	12.42	2.887	93
			Total	12.42	2.887	93
		Total	LOWCoP	7.67	3.055	3
			HighCoP	11.68	2.798	146
			Total	11.60	2.849	149
Total	Low	Low	LOWCoP	3.29	2.453	171
			Total	3.29	2.453	171
		High	LOWCoP	8.00	.	1
			HighCoP	9.50	.707	2
			Total	9.00	1.000	3
		Total	LOWCoP	3.31	2.472	172
			HighCoP	9.50	.707	2
			Total	3.39	2.546	174
	High	Low	LOWCoP	7.08	1.977	13
			HighCoP	9.92	2.186	80
			Total	9.53	2.366	93
		High	HighCoP	12.18	2.973	103
			Total	12.18	2.973	103

	Total	LOWCoP	7.08	1.977	13
		HighCoP	11.20	2.879	183
		Total	10.92	3.006	196
Total	Low	LOWCoP	3.55	2.607	184
		HighCoP	9.92	2.186	80
		Total	5.48	3.843	264
	High	LOWCoP	8.00	.	1
		HighCoP	12.13	2.968	105
		Total	12.09	2.981	106
	Total	LOWCoP	3.58	2.620	185
		HighCoP	11.18	2.870	185
		Total	7.38	4.691	370

Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.721	184.672[b]	5.000	357.000	.000	.721
	Wilks' Lambda	.279	184.672[b]	5.000	357.000	.000	.721
	Hotelling's Trace	2.586	184.672[b]	5.000	357.000	.000	.721
	Roy's Largest Root	2.586	184.672[b]	5.000	357.000	.000	.721
K_Total_2_groups	Pillai's Trace	.153	12.925[b]	5.000	357.000	.000	.153
	Wilks' Lambda	.847	12.925[b]	5.000	357.000	.000	.153
	Hotelling's Trace	.181	12.925[b]	5.000	357.000	.000	.153
	Roy's Largest Root	.181	12.925[b]	5.000	357.000	.000	.153
C_Total_2_groups	Pillai's Trace	.123	10.025[b]	5.000	357.000	.000	.123
	Wilks' Lambda	.877	10.025[b]	5.000	357.000	.000	.123
	Hotelling's Trace	.140	10.025[b]	5.000	357.000	.000	.123
	Roy's Largest Root	.140	10.025[b]	5.000	357.000	.000	.123
R_Total_2_groups	Pillai's Trace	.073	5.633[b]	5.000	357.000	.000	.073
	Wilks' Lambda	.927	5.633[b]	5.000	357.000	.000	.073
	Hotelling's Trace	.079	5.633[b]	5.000	357.000	.000	.073
	Roy's Largest Root	.079	5.633[b]	5.000	357.000	.000	.073
CategoryofPerformance	Pillai's Trace	.033	2.414[b]	5.000	357.000	.036	.033
	Wilks' Lambda	.967	2.414[b]	5.000	357.000	.036	.033
	Hotelling's Trace	.034	2.414[b]	5.000	357.000	.036	.033
	Roy's Largest Root	.034	2.414[b]	5.000	357.000	.036	.033
K_Total_2_groups * C_Total_2_groups	Pillai's Trace	.000	.[b]	.000	.000	.	.
	Wilks' Lambda	1.000	.[b]	.000	359.000	.	.
	Hotelling's Trace	.000	.[b]	.000	2.000	.	.
	Roy's Largest Root	.000	.000[b]	5.000	356.000	1.000	.000
K_Total_2_groups * R_Total_2_groups	Pillai's Trace	.047	3.500[b]	5.000	357.000	.004	.047
	Wilks' Lambda	.953	3.500[b]	5.000	357.000	.004	.047
	Hotelling's Trace	.049	3.500[b]	5.000	357.000	.004	.047
	Roy's Largest Root	.049	3.500[b]	5.000	357.000	.004	.047
K_Total_2_groups *	Pillai's Trace	.000	.[b]	.000	.000	.	.
	Wilks' Lambda	1.000	.[b]	.000	359.000	.	.
	Hotelling's Trace	.000	.[b]	.000	2.000	.	.

CategoryofPerformance	Roy's Largest Root	.000	.000[b]	5.000	356.000	1.000	.000
C_Total_2_groups * R_Total_2_groups	Pillai's Trace	.000	.[b]	.000	.000	.	.
	Wilks' Lambda	1.000	.[b]	.000	359.000	.	.
	Hotelling's Trace	.000	.[b]	.000	2.000	.	.
	Roy's Largest Root	.000	.000[b]	5.000	356.000	1.000	.000
C_Total_2_groups * CategoryofPerformance	Pillai's Trace	.000	.[b]	.000	.000	.	.
	Wilks' Lambda	1.000	.[b]	.000	359.000	.	.
	Hotelling's Trace	.000	.[b]	.000	2.000	.	.
	Roy's Largest Root	.000	.000[b]	5.000	356.000	1.000	.000
R_Total_2_groups * CategoryofPerformance	Pillai's Trace	.000	.[b]	.000	.000	.	.
	Wilks' Lambda	1.000	.[b]	.000	359.000	.	.
	Hotelling's Trace	.000	.[b]	.000	2.000	.	.
	Roy's Largest Root	.000	.000[b]	5.000	356.000	1.000	.000
K_Total_2_groups * C_Total_2_groups * R_Total_2_groups	Pillai's Trace	.000	.[b]	.000	.000	.	.
	Wilks' Lambda	1.000	.[b]	.000	359.000	.	.
	Hotelling's Trace	.000	.[b]	.000	2.000	.	.
	Roy's Largest Root	.000	.000[b]	5.000	356.000	1.000	.000
K_Total_2_groups * C_Total_2_groups * CategoryofPerformance	Pillai's Trace	.000	.[b]	.000	.000	.	.
	Wilks' Lambda	1.000	.[b]	.000	359.000	.	.
	Hotelling's Trace	.000	.[b]	.000	2.000	.	.
	Roy's Largest Root	.000	.000[b]	5.000	356.000	1.000	.000
K_Total_2_groups * R_Total_2_groups * CategoryofPerformance	Pillai's Trace	.000	.[b]	.000	.000	.	.
	Wilks' Lambda	1.000	.[b]	.000	359.000	.	.
	Hotelling's Trace	.000	.[b]	.000	2.000	.	.
	Roy's Largest Root	.000	.000[b]	5.000	356.000	1.000	.000
C_Total_2_groups * R_Total_2_groups * CategoryofPerformance	Pillai's Trace	.000	.[b]	.000	.000	.	.
	Wilks' Lambda	1.000	.[b]	.000	359.000	.	.
	Hotelling's Trace	.000	.[b]	.000	2.000	.	.
	Roy's Largest Root	.000	.000[b]	5.000	356.000	1.000	.000

K_Total_2_groups * C_Total_2_groups * R_Total_2_groups * CategoryofPerformance	Pillai's Trace	.000	.[b]	.000	.000	.	.
	Wilks' Lambda	1.000	.[b]	.000	359.000	.	.
	Hotelling's Trace	.000	.[b]	.000	2.000	.	.
	Roy's Largest Root	.000	.000[b]	5.000	356.000	1.000	.000

a. Design: Intercept + K_Total_2_groups + C_Total_2_groups + R_Total_2_groups + CategoryofPerformance + K_Total_2_groups * C_Total_2_groups + K_Total_2_groups * R_Total_2_groups + K_Total_2_groups * CategoryofPerformance + C_Total_2_groups * R_Total_2_groups + C_Total_2_groups * CategoryofPerformance + R_Total_2_groups * CategoryofPerformance + K_Total_2_groups * C_Total_2_groups * R_Total_2_groups + K_Total_2_groups * C_Total_2_groups * CategoryofPerformance + K_Total_2_groups * R_Total_2_groups * CategoryofPerformance + C_Total_2_groups * R_Total_2_groups * CategoryofPerformance + K_Total_2_groups * C_Total_2_groups * R_Total_2_groups * CategoryofPerformance

b. Exact statistic