Title:
Standardized testing in the Caribbean: exploring the relationship between high stakes and low stakes testing in secondary schools in Trinidad and Tobago.

Author:
Melanie-Marie Haywood
University of Trinidad and Tobago
O’MEARA CAMPUS
Abstract

This study explored the relationship between National Certificate of Secondary Education (NCSE) level one, and Caribbean Secondary Education Certificate (CSEC) in the subjects of Mathematics and Language Arts/English. These two assessments are two major standardised assessments carried out at a national and regional scale respectively. The study sampled test results for students in 16 schools between 2007-2015. Corresponding results were used for students who wrote examinations between 2009-2017. For the qualitative component of the study, the researcher sampled 14 teachers for focus groups.

Findings showed that relationship exists between NCSE level one and CSEC in Mathematics and English/Language Arts respectively, and overall buy-in for NCSE is significantly lower than that of CSEC. The study also found that attitude toward schooling is a complex variable that does not necessarily presuppose improved academic performance. Statistical analysis revealed that gender and type of school are two important variables affecting of academic performance. Findings also revealed a rank in the performance of schools where single-sex denominational schools are highest performers, and government schools are lowest performers.

Keywords. NCSE, CSEC, standardised testing, low-stakes, high-stakes, attitude, academic performance.
1.0. Introduction

The purpose of education is to equip and assist the student in going from the unknown to the known. The nature of the education system is that it develops in stages, and the idea is that at each stage, the learner is able to grasp the concepts attached because moving on to another stage would be impossible if a true understanding of the previous stage is not attained (Erikson, 1968). The ability to evaluate and assess academic success has become increasingly important as can be seen from the growing literature on assessment in schools, and we can see globally that there are a significant number of learners who are not grasping the concepts necessary at each stage, resulting in failure or drop outs (Edwards, 2007). Therefore, it is possible to assume that the enrolment, success, and even sometimes the perseverance of students are measures used to determine the success of an educational system. A type of evaluation of academic success comes through standardized testing.

In the past decade, a number of changes within the education system in Trinidad and Tobago have raised concern and alarm amongst key stakeholders. One of the biggest issues in education reform has been the development and frequency of standardized testing of children from the beginning of primary school (National Test, and Secondary Entrance Assessment (SEA)) to the end of secondary school (National Certificate of Secondary Education (NCSE), and Caribbean Secondary Education Certificate (CSEC)). Teachers, parents, and even students complain that the rigidity and requirements of meeting testing deadlines is causing a significant barrier to the freedom of learning (Edwards, 2007). Teachers being forced to complete curriculum in a particular time frame, and students being forced to learn within this same time results in great frustration for both groups. The purpose of the frequency of these mandatory tests are, to date, still being called into question. However, since the testing is mandatory educators must do what is required to prepare students for examinations.

The Caribbean Examinations Council (CXC) was instituted in 1972 when the Caribbean Community and Common Market (CARICOM) saw fit to institute a standardized form of testing and certification that would be accepted throughout the region. (CXC, 2018) CXC represents 16 English speaking countries and territories in the region. CXC has completely replaced the General Certificate
of Education (GCE), which was the original European certification used throughout the region at the end of secondary school. GCE has been replaced by CSEC which students now take at the end of secondary school as their certification to be accepted into further education.

The NCSE, however, is a national certification for Trinidad and Tobago and was introduced in 2005 to secondary school students. This exam was to be used as a form of continuous assessment for secondary school students and includes both a written examination component (level one) offered at the end of Form 3, and a continuous project component (level two) that is completed throughout Forms IV and V. At level one, students are usually between age 13-14 years old when writing the examination.

According to Roberts (1968), “educational attainment is one of the important contributing factors toward career choice” (p. 176). In Trinidad and Tobago universities base their acceptance criteria for high school leavers, on their qualifications attained either at CSEC or CAPE. The NCSE level one examination is the first national standardized assessment that all high school students complete in Trinidad and Tobago, but the purpose of this assessment is called into question. It is true that national assessment and continuous assessment can be conducted by any governing education board to measure the success of the system, but the amount of resources utilised in the administration of this exam should be a measure of the utility of the data gathered after the test is administered. However, in Trinidad and Tobago, NCSE is completed at the end of the second term of Form 3. Data from the test is gathered at the end of the year, and therefore it is not useful to students, parents, or educators to determine placement in Form 4 where students begin preparation for CSEC examinations to be completed in the following year.

1.1. Research Questions

1. What is the relationship between NCSE Level 1 and CSEC performance for secondary school students in Trinidad and Tobago?

2. What factors in a student’s experience between NCSE Level 1 and CSEC impact high school student’s attitude toward schooling in Trinidad and Tobago?
2.0. Theoretical approaches

The standardized testing evaluated in this study is a form of assessment that measures a point-in-time level of knowledge acquisition as it is a point-in-time examination. The progress of knowledge acquisition is taken into consideration by the continuous assessment component of both examinations. As such, the theories of learning discussed in this study come from a combination of a number of educational theorists who posit their observations and expectations of the learner, and how one’s interaction impacts the achievement of learning goals.

Philippe Parmentier’s theory of learning. Parmentier’s (1996) theory of learning was developed based on a study that explored the notion that a student’s personal history, attitude toward schooling, and perception of schooling all impact overall academic performance as measured by testing. Vandamme, Meskens, and Superby (2007) go on to elaborate by explaining that a “student’s intermediate and final academic performance is influenced by the interaction of [the above-mentioned] three sets of factors” (p. 10).

Parmentier (1996) describes a student’s personal history as a stable variable that includes an individual’s basic demographic information such as age, socio-economic background, gender etc. A student’s attitude toward schooling is a process, or changing factor. This, he believes, is measured by a student’s participation, or willingness to participate in school related activities – both curricular and extra-curricular (Dreyakulu as cited in Arslantas, 2016). Green et.al. (as cited in Arefi, Naghibzadeh, & Boloki, 2014) have defined a student’s perception of schooling by showing that “Research has also supported the view that academic achievement and academic self-concept mutually reinforce each other, to the extent that a positive or a negative change in one facilitates a commensurate change in the other... Put simply, improved academic self-concepts will lead to better academic achievement, and improved achievement will lead to better academic self-concepts.” (p. 1).

Piaget developed a theory of learning that spans the lifetime of the student. The purpose of this theory is to find explanation and meaning in the behaviours of the learner throughout the various stages of their existence. The reason for identifying the various stages, according to Piaget (1954), is
that what may be apparent to an adult, may not be as apparent to a child. Therefore, the stages of cognitive development allow for new types and levels of understanding in the learner.

Bronfenbrenner (1989) took Erikson’s appreciation of the social context of the learner to a deeper level of penetration on cognitive development. The theory essentially purports that there are numerous factors that impact the development of the learner, and these should be effectively contextualized in all learning experiences. Bronfenbrenner (1989) goes on to name some of the most important social contexts that can significantly impact the development of the human:

1. The family – parenting styles, culture, divorce (Bronfenbrenner & Evans, 2000)
2. Peers – aggression, acceptance, bullying
3. Teachers – caring, support, and abuse (Woolfolk Hoy, & Weinstein, 2006; Harris, 1998)

In the context of assessment in learning, it is imperative that these factors be considered as the learning experience can be extremely skewed by changes in these areas.

3.0. Standardized testing

Standardized testing in education is a form of assessment that when administered, is “done so according to certain rules and specifications so that testing conditions are the same for all test takers” (Johnson-Centre, n.d., para. 1). Standardized tests come in many forms; however, for the purpose of this paper as pertains to NCSE and CSEC, the focus is on the written assessment. For the purpose of this paper, the researcher used Wang, Beckett, and Brown’s (2006) definition of a standardized test:

- a large scale, externally developed and mandated, uniformly administered and scored evaluation of student learning . . . (a) is externally imposed by the state government; (b) assesses state-prescribed content standards; (c) follows a uniform procedure in administering, scoring, and interpreting the test; and (d) the results are often used to determine rewards and sanctions for students, teachers, schools, or districts. (p. 306)

The purpose of standardizing is generally to provide a form of assessment that can be used to evaluate performance across a large body of individuals (Wang, Beckett, & Brown, 2006). This
allows for comparison, and even understanding in evaluating the progress or status of a body of students. Concerning the main use of standardized testing, according to Holmes (2010), “It is used to determine student achievement, growth, and progress” (para. 1). However, the questions of validity and reliability are always at the forefront of any type of research or measurement; how sure can we be that measure is a determinant of achievement, growth, or progress?

The most globally known standardized test today is the Scholastic Aptitude Test (SAT), and it is one of the longest standing standardized assessments. In 1926, the SAT was founded by a non-profit group of educational organizations as one 90-minute test, but grew into what we know it to be now, a math and a verbal examination. By the end of World War II, the test was accepted by most universities in the USA as a selection and entrance criteria, and has remained mostly unchanged.

3.1. Types of standardized testing

The three major categories of standardized testing fall into either high-stakes, medium-stakes, or low-stakes testing. High-stakes testing is the test administered that has direct personal impact on the examinee’s life and goals (Barry, Horst, Finney, Brown, & Kopp, 2010). Examples of this type of test are the SAT, or in the case of Trinidad and Tobago, and the Caribbean, the CSEC exam. Low-stakes testing on the other hand is opposite to high-stakes where the exam has very little impact on the individual examinee’s goals and personal development (Barry et al., 2010). Examples of this are the PIRLS international exam on reading and literacy, and in the case of Trinidad and Tobago, the NCSE level 1 examination. Medium-stakes testing falls in between these two categories where the value to the students and implication of the exam are higher than in low stakes, but not as high as in high-stakes tests (Barry et al., 2010). Essentially, the consequences of the test often determine the level of the stakes.

3.1.1. High-stakes testing. High-stakes testing, according to Nichols, Glass, and Berliner (2012), is “the process of attaching significant consequences to standardized test performance with the goal of incentivizing teacher effectiveness and student achievement” (p. 3). The concept of high-stakes is that by making the consequences of the results attached to either a strongly positive, or
strongly negative real life outcome the test becomes a significant milestone in the lives of both
students and teachers (Nichols, Glass, & Berliner, 2012). This then makes the test a factor that is no
longer affected solely in the classroom, but is affected by factors that exist outside of the classrooms
walls. Essentially, the high stakes associated with testing form a strong motivator to both students and
teachers in the performance on the test, and preparation for the test.

### 3.1.2. Low-stakes testing

According to Sessoms and Finney (2015), low-stakes tests are
common, and are used to provide evidence of learning and not much else. Although these tests tend to
have no personal consequences to the student, they are still high in stakes for the institutions that
choose to administer them as they result in comparisons to the other local institutions, and in some
cases international rankings for quality and success in education.

Research has shown that there is a value to low-stakes testing. According to Schrank (2016), frequent low-stakes testing is a “more effective way to boost student learning and knowledge
retention over time compared to infrequent high-stakes testing” (p. 119). This has been found to be
the case in students who are being exposed to new content, topics, subjects, and even levels of
education (Shrank, 2016). This is conditional on the fact that frequent low-stakes testing allows the
student to exercise the skill of new ways of thinking, which requires practice and retrieval skills. The
value in low-stakes testing can also be seen in the fact that it forms a sense of continuous assessment
for the student when feedback is timely and comprehensive (Pennebaker et. al., 2013). A study
conducted by Pennebaker et al. (2013) found that low-stakes testing is highly valuable in the overall
development of the educational outcomes of the students. Findings showed that the experimental
group performed better on the quizzes, and retained more course information at the end of the term
compared to the control group. Correlation was also found between the frequency of testing and the
reduction of the performance scoring gap of students with low and high socioeconomic backgrounds.

The major point of concern and research in connection to low-stakes testing has become the
concept of examinee effort, also known as student motivation (Steedle, 2014; Zilberberg, Finney,
Marsh, & Anderson, 2014). This is an international matter that spans major countries and education
systems such as the USA, Canada, Saudi Arabia, Sweden, and the Caribbean. As a result of this,
questions of validity arise in the analysis of low-stakes testing considering the impact of examinee effort in the scheme of assessment.

### 3.1.3. Examinee effort/student motivation

Concerns with examinee effort are strongly associated with low-stakes testing in educational research. This is because the issue of motivation for success in low-stakes testing often occurs, raising questions of relevancy and validity of the low-stakes examination (Barry et al., 2010). Examinee effort in high-stakes tests is likely to be high as a result of the direct personal impact the success or failure of the examination has on the individual. Examinee effort in low-stakes assessments tends to be much more variable based on the limited implications of failure (Barry et al., 2010).

The implications of test-taking effort for the student’s test results have been found to be significant, and as such, in the US, a number of agencies including the American Educational Research Association, and the American Psychological Association have recommended that test-taking effort be collected and reported along with test scores for accuracy in interpretation (American Educational Research Association, & American Psychological Association, 1999). Also, the International Test Commission (ITC) has made the call for those administering national tests “to consider other qualities which may have artificially lowered or raised results when interpreting scores” (International Test Commission, 2000, p. 15).

Research on the impact of test-taking motivation on the performance in the examination, has found that students who are more motivated consistently out motivate students who have lower motivation (Wise & DeMars, 2005). The low motivation has been found to even be able to obscure the presence of high ability – rendering invalid the test’s ability to correctly reflect the ability of the student (Wise & DeMars, 2005). Essentially, the validity of the assessment, according to educational researchers, cannot be possibly established without the measurement of test-taking effort (Mathers, Finney, & Hathcoat, 2018).
4.0. Education Reform in Trinidad and Tobago

Education reform in small states has been mostly done in connection with internal funding agencies such as the World Bank, and the Inter-American Development Bank (De Lisle, 2012). This often causes transfer of intentions from the donor countries into the small states, leaving the intentions and needs of the small states unattended to (Ibrahim, 2010).

According to De Lisle (2012), since its independence in 1962, Trinidad and Tobago has gone through three reform stages, and is in its fourth stage at present:

1. The 15-year Education Plan (1968-1983)

Secondary Education Modernization Plan (SEMP). Under SEMP, education reform in Trinidad and Tobago took a systematic approach focusing on the primary school sector first, and then moving on to the secondary school sector. The purpose of the SEMP initiative was to improve the quality of education in the secondary school system (IDB, 2009). Although in 2009, the Seamless Education Reform began, it was not to the neglect of the secondary school system, where objectives within this reform still focused on the improvement in both primary and secondary school systems.

The implementation of SEMP was a whole-system education reform. This type of reform is also known as a large-scale system reform, which comes with its own challenges for success (De Lisle, 2012). Particularly in the case of the SEMP program, much of the change initiative was still superficial at the end of its tenure, “with mechanisms and structures for sustainability notably absent” (De Lisle, 2012, p. 67).

According to Atchoarena, Dias da Graça, and Manuel Marquez (2008), the improvement and development of the secondary school system is vital to the development of the small state as this system prepares the pathway for the workforce, and the quality of education through to tertiary education. Despite how vital secondary education reform has been, researchers still argue that small
states face a large challenge with its implementation (Di Gropello, 2006; De Lisle, 2012). The reason for this has been posited by Holsinger and Cowell (2000), where the difficulty is often attributed to the colonial ideology of the state and education still being pervasive in society, making implementation of reform that varies from this ideology a challenge. In Trinidad and Tobago, to monitor the progress of education reform, a number of standardized tests were implemented or adapted at the primary and secondary school levels.

Education reform in Trinidad and Tobago, is also met with disdain and confusion from the key stakeholders involved (Joseph, 2010). The reform strategies are often perceived to be handed down from external agencies, and do not fit the demographic they intend to change (De Lisle, 2012). As such, there is often less government control and strategic choice, leaving a change strategy that is “borrowed, distorted, or non-existent” (De Lisle, 2012, p. 67). Added to this issue is the fact that there is extremely limited data on the culture and history of the country for effective analysis of needs in the education sector, which often creates a vacuum between the strategy, and the actual implementation of the plan (Louisy, 2004).

De Lisle (2012) in his comprehensive review of the SEMP program was able to provide findings to state that the program essentially was not able to achieve its goals. Barriers to change were numerous, and existed in various areas of the strategy and its implementation. The first, and most frequent was that of the ambiguity in the reform goals even in the foundation of the strategy, and then there was clear ambiguity in the communication of these goals to the key stakeholders. As such, it was found that teachers and other ‘ground’ workers did not buy into the program, and resisted it strongly. Other barriers included the obvious confusion of teachers and other ‘ground’ workers as to the purposes of many of the changes, and lack of training necessary for implementation of these changes.

The lack of communication created an ever-greater problem for the SEMP program; De Lisle (2012) identified that there was a significant disconnect between the roles of individuals involved in the strategy, the ideas of those who were involved, and the institutions that they were to represent. The NCSE examination came out of the SEMP program and was identified as a weak area in this
barrier to change, in that the NCSE program was implemented across divisions in the MOE, but there was not always cohesion between the units (De Lisle, 2012).

Aside from the human aspect of implementation, the management of resources, both human and physical, was identified as ineffective. There were a number of projects occurring under the SEMP program at once, and the resources were strained and often times did not have the necessary training to effectively implement the projects (De Lisle, 2012).

Education professionals are referring to the burst in testing as a “repression of ideas in our education system” (Edwards, 2007, para. 1). Even after the implementation of so many tests, the then Minister of Education confirmed that only “30% of high school graduates achieve a full CXC certificate” (Spence, 2011, para. 3). In secondary schools, there are only two mandatory standardized tests administered – NCSE level 1 which is a low-stakes test, and CSEC, which is a high-stakes test.

4.1. National certificate of secondary education. NCSE level 1 was introduced into Trinidad and Tobago in 2006. It is administered at the end of Form 3 as a form of continuous assessment for government assisted, and public high schools. The purpose of the examination is to determine that students have achieved learning outcomes in the following subjects: language arts, mathematics, science, social studies, visual and performing arts, Spanish, physical education, and technology education. Level 2 of NCSE is determined by the work students complete in Forms 4 and 5.

According to the Secondary Education Modernization Programme (SEMP) (2011), NCSE seems to be proving useful in indicating growth in academic performance of high school students:

Improved student performance in the National Certificate of Secondary Education (NCSE) exam; more specifically, a 16.8 percentage point increase in the pass rate between 2006 and 2009. Two of the highest increases were in Science and Math, 108% and 67% respectively. (p. 5)

However, according to a study done to determine the perspective of teachers, parents, and students of the NCSE Level 1, Hislop (2012) was able to conclude that “doubt still exists in the mind
of some parents, teachers, and students about the purpose of the examination, whether it is really important and the use to which the NCSE Level 1 certification can be put” (p. 50).

4.2. Caribbean secondary education certificate. The certification that is considered the most significant in the secondary school system, and understandably so, is CSEC. This examination is administered at the end of Form 5, and is not a national examination, but is a regionally administered examination coming from a regional examination body known as the Caribbean Examinations Council (CXC). There are 33 subjects offered at this level – 28 of which are General Proficiency, and five of which are Technical Proficiency (Caribbean Examinations Council, 2011). Most universities in TT accept students with five CSEC subjects, and most entry level jobs in the workplace require five CSEC subjects as well. Achieving grades I, II, or III, in five CSEC subjects is considered a complete CXC certificate (CXC, 2018).

The general consensus in Trinidad and Tobago is that CSEC is essential. Any country must have some sort of marker to determine the academic progress of a student through the compulsory education system, and this marker should be standardized. However, there has been recent uproar that this assessment doesn’t cater for the various learning styles and professional preferences of all students in the system. In 2012, there was still a 60% failure rate for students obtaining a full CXC certificate (Trinidad and Tobago Newsday, 2012). As such, CXC created a curriculum and examination for technical subjects.

Citizens of Trinidad and Tobago still seem to believe that the education system is failing their youth. Blaming the criminality in society on the failure of the education system to do its part is not a difficult inference to make with numbers such as 60% failure rates at CSEC (Spence, 2011).

5.0. Methods

As a mixed methods study, the nature of the data collected was in both quantitative, and qualitative forms. The concurrent triangulation design allows for a researcher to gather both quantitative and qualitative types of data, and integrate the results after the analysis phase.
5.1. Sample

The study used national NCSE Level One results of students who wrote the examination between 2007-2015 in 16 schools sampled from each education district in Trinidad and Tobago. Corresponding CSEC results from 2009-2017 were used for comparison. The study uses a sample of all types of schools in Trinidad and Tobago that take the NCSE examination. This consists of both government assisted (private, denominational), and government secondary schools. Furthermore, the researcher sampled 14 teachers who have taught standardized exam preparation for at least two years. This means that the teachers have taught either Form 3 (ages 13/14) or Form 5 (ages 15/16) for at least two years during their teaching career. At the time of data collection, all teachers were registered teachers with the Ministry of Education, working in a secondary school teaching either Form 3 or Form 5, or a principal in order to meet the criteria for sampling.

5.2. Instruments

There were three data collection methods used in this study – survey, test results, and focus groups. The survey and test scores were used as data collection instruments for secondary school students. The test scores for 14 schools were attained directly from the MOE’s DERE. The survey was administered face-to-face in a purposive selection of schools throughout Trinidad and Tobago. All schools sampled (for both test results, and the survey) were selected to be representative of the following categories:

1. Denominational
2. Government
3. Government Assisted
4. Single-sex (all-girls/all-boys)

5.2.1. Test Results. Between 2008 and 2017, 143 schools registered students to participate in the NCSE level 1 examination, and 162 schools registered students to participate in the CSEC examinations. As such, a sample of 14 schools geographically distributed to represent each
educational district (two schools from each district) in Trinidad and Tobago were selected. NCSE Level 1 and CSEC test results were provided by the MOE’s DERE.

5.2.2. Survey. The survey for this study was a self-designed survey created in consultation with a number of accepted and established surveys used in education and psychological research based on the theoretical approaches. The survey is cross sectional: evaluating the students at one point in time. As a result of the nature of the data to be found, the survey items are in multiple choice, dichotomous, and Likert format. The survey was administered to 400 participants, with 364 surveys returned attempted, and of this 364 attempted 202 surveys returned completed.

5.2.3. Focus group. The researcher conducted two focus groups with a total of 14 teachers (six teachers from schools in North Trinidad, and eight teachers from schools in South Trinidad). In order to qualify, participants had to have taught either NCSE Level 1 or CSEC preparation for a minimum of two years. Teachers for the focus group were sampled through snowball sampling. The duration of the focus group sessions ranged from 60-90 minutes.

6.0. Findings

In this study, the core phenomenon would be the existence, or lack thereof, of a statistically significant relationship between the two test results. Results between both exams were standardized and analysed to determine a relationship. Upon establishing the statistical relationship, the categories surrounding the core phenomenon were discussed by the results of the survey and focus group data.

6.1. Pilot study

Initially, the survey was administered to a pilot group in order to determine the reliability of its scales. The pilot was conducted in one school on 33 students, before the major data collection exercise began. This process took one week from attaining approval to the collection of the completed pilot surveys. Cronbach’s Alpha was used to determine the reliability of the scale. The result of the analysis revealed an alpha value of 0.82442 which shows a high level of internal consistency for the survey.
6.2. Statistical analysis.

Frequency testing was done for confirmatory purposes, but also to assist in the descriptive analysis of the data. Frequencies were conducted on test results on each variable, for each year of test results. Comparisons between the NCSE Level 1 and CSEC were conducted in each subject.

Upon the completion of establishing the scale reliability and validity, the survey data, and the test results for both NCSE level 1 and CSEC were tested for normality using the Kolmogorov-Smirnov test. The significance statistic in the data shows that the data set was not normally distributed, and as such, parametric testing could not be conducted in order to draw inferences from the statistical data. As such, the non-parametric test alternatives were used.

Firstly, the chi-square tests for association were conducted in order to show associations between the variables, and the results of the NCSE and CSEC examinations. The results showed that there were associations between the variables measured, and as such, the study went on to conduct stronger tests in order to answer the hypotheses.

The researcher used the Mann Whitney U test to compare the medians of the results of the NCSE level 1 and CSEC results to determine statistical significance in the relationship between the two variables. Further to this, Spearman’s R was used to establish correlations within each exam (by year) to determine if there was a relationship between the variables of gender, school, and year, and the performance of the student in that examination category. These findings were supported by the chi-square tests conducted earlier on. All comparative analyses of variables through frequencies were conducted with standardized results.
6.3. Focus Groups

Table 1.

Participant descriptions

<table>
<thead>
<tr>
<th>Location</th>
<th>Participant</th>
<th>Type of Schools</th>
<th>No. of Years’ Experience</th>
<th>Subjects taught</th>
</tr>
</thead>
<tbody>
<tr>
<td>North Trinidad</td>
<td>B</td>
<td>Government, Private</td>
<td>43</td>
<td>Music, English, Religious Education</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>Denominational &amp; Gov’t</td>
<td>16</td>
<td>Mathematics, Social Studies</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>Government</td>
<td>9</td>
<td>All Sciences, Mathematics</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>Government</td>
<td>8</td>
<td>ICT, Mathematics</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>Government</td>
<td>29</td>
<td>English</td>
</tr>
<tr>
<td>South Trinidad</td>
<td>G</td>
<td>Private, Government</td>
<td>9</td>
<td>English</td>
</tr>
<tr>
<td></td>
<td>H</td>
<td>Denominational &amp; Gov’t</td>
<td>18</td>
<td>Music, Theatre Arts</td>
</tr>
<tr>
<td></td>
<td>I</td>
<td>Denominational</td>
<td>7</td>
<td>OA, Accounts, Economics, Business</td>
</tr>
<tr>
<td></td>
<td>J</td>
<td>Denominational &amp; Gov’t</td>
<td>23</td>
<td>Mathematics, Accounts</td>
</tr>
<tr>
<td></td>
<td>K</td>
<td>Denominational</td>
<td>19</td>
<td>Biology, Chemistry, HSB</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>Denominational</td>
<td>16</td>
<td>Physical Education and Sport</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>Denominational &amp; Gov’t</td>
<td>12</td>
<td>Geography</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>Government</td>
<td>8</td>
<td>English</td>
</tr>
</tbody>
</table>

6.3.1. Challenges with NCSE examination preparation. All of the teachers vocalized that
their greatest challenge with the NCSE examination is that the students do not take it seriously. Some
teachers presumed that this was as a result of the lack of information the students and teachers get on
the exam’s value, as well as the fact that teachers also do not see the value of the examination and this
transfers to the students. All teachers unanimously agreed that the examination is extremely
undervalued by students, and as such, they do not take preparation, nor results seriously.

6.3.2. The value of NCSE. The value of NCSE was the major question raised by all teachers
when asked anything about the examination. The consensus among all of the teachers was that the
quality of the examination has changed – worsened over the years. This worsening includes not only
the exam setting, but the marking of the exam, as some teachers raised questions on the authenticity of
the marking of the exam. It was noted that students who did poorly in NCSE examinations, thrived in
the same subject with internal exams, and even with CSEC results. Teachers confirmed that schools
maintain their own internal examinations for continuous assessment purposes, and that the NCSE
poses very little effect on the schools’ measurements.

Teachers also agreed that the exam’s greatest value could be found in the timing, but loses its
value as a result of its current timing in the school term. The exam being placed at the end of Form 3,
with results being awarded in Form 4 proved to be “very useless” to teachers, students, and parents.
Teachers indicated that the exam would be much more valuable if placed earlier on in the students’
schooling, as well as students taking all of the subjects from an earlier stage in their education. All
teachers were of the opinion that the preparation for CSEC was a much easier process on the part of
the students. This in no way minimized the consistent feedback that there were still major issues with
preparation for CSEC.

6.3.3. Ways to improve exam preparation and performance. Teachers presented several
methods they believed could impact positively on student performance in the CSEC examination:

1. Parental involvement
2. Clearer communication from the MOE
3. More resources
4. Collaborative teaching
5. More counselling and career guidance
6. More time to teach subject matter and less time on past papers
7. Preparing students for CSEC choices from entrance into secondary school

6.3.4. Comparing the performance between NCSE and CSEC. Teachers were of the
unanimous opinion that there is a difference in preparing for the two examinations, both as a teacher
and as a student. Furthermore, the teachers believed that students take the CSEC examination much
more seriously, and as a result, preparation is easier for the teachers. The lack of information that
teachers receive concerning the value and even assessment methods for NCSE make it much more
difficult to prepare as a teacher, and therefore prepare the students and parents.
6.4. Test Scores and Survey findings

There were 7,557 students who took the Language Arts examination, and the same number of students sat the Mathematics examination in the sample. The majority of students came from government schools, and the majority of students who took the exam were male students. Also, the average result for both Language Arts and Maths is the grade D which is a failing grade.

*Figure 1.* Line graph showing distribution of students by grade attained in Language Arts (NCSE).

*Figure 2.* Line graph showing distribution of students by grade attained in Mathematics (NCSE).
More students passed the Language Arts examination (cumulative percentage of 32.8), than in the Mathematics examination (cumulative percentage of 22.7). In both exams, we can see a sharp peak in the number of students that have failed the exam, or did not receive a mark for it – the peak is sharper in Mathematics than in Language Arts. Overall, more students fail this exam with an average passing mark of D.

A total of 17,439 English A and Mathematics examinations were written at the CSEC level – due to data protection, the MOE was not able to release the total number of students in the sample, simply the results per paper. Similar to that of the NCSE results, the majority of students were male, and the schools and therefore types of school also remained the same, as these variables are unchanging due to the sampling methodology used. We can also see from the Mode above that more students overall wrote Mathematics, than wrote English A. Overall, the majority of students have passed CSEC Mathematics and English A in the past five years, with a cumulative percent of 55% attaining a grade I-III. As a result of the sample being the same for both NCSE and CSEC, the frequency distributions for school, and type of school do not differ.

![Figure 3. Percentages of students by grade for CSEC.](image)
Figure 4. Frequency distribution of students grades in CSEC English A.

Figure 5. Line graph representing frequency distribution of grades attained in CSEC Mathematics.
On average, frequency statistics show that students who take the English A CSEC exam attain a grade III, while the most commonly attained grade is a grade I. On average, statistics show that the students that take Mathematics CSEC attain a grade IV, while the most commonly occurring grade is V.

6.4.1. Comparison of NCSE and CSEC results. The results of the two examinations are analysed by subject, looking at trends in the results of these exams by corresponding years of NCSE and CSEC where the students who took the NCSE exam in 2011 are compared with CSEC results from 2013 over a five-year period. Consistently, in both math and language art, the pattern of results is the same between NCSE and CSEC:

1. In higher passing grades (A/I-B/II) results show a clear inverse relationship with higher percentages of students passing at CSEC and the inverse amount with significantly lower percentages passing at NCSE. More students pass CSEC with higher grades.

2. In lowest passing grade C/III, the results are almost the same between NCSE and CSEC, where both subjects have similar percentages of students passing/failing at similar stages. A similar percentage of students who pass with average grades at NCSE pass with average grades at CSEC.

3. In failing grades, results show a direct relationship, with significantly more students failing at NCSE than CSEC. This shows that overall, more students pass CSEC than NCSE.

6.5. Hypothesis testing

\( H_1 \) There is no statistically significant difference between NCSE Level 1 Mathematics and CSEC Mathematics results.

\( H_2 \) There is no statistically significant difference between NCSE level 1 Language Arts and CSEC English A results

With a significance level of 0.00 the null hypotheses were rejected, showing that there is a statistically significant difference between the two examinations in both English and Mathematics. This is confirmatory of what has been found using frequency and chi-square testing for association
that found that often times the relationship between the results of the two tests are either positive, or negative.

\( H_3 \) There is no statistically significant relationship between personal history, and attitude toward schooling

Results showed that there was not enough evidence to reject the null hypothesis. However, upon further investigation, the researcher was able to still find a statistically significant relationship between Gender, Type of School, and Attitude toward schooling.

The results of Spearman’s Rho show that there is a statistically significant relationship between Type of school and Attitude toward schooling (sig. 0.026); Gender and Attitude toward schooling (sig. 0.002). Interestingly, statistics also indicate that the attitude toward schooling of the students in the various types of school does not directly relate to the performance of the school.

Findings show the following rank in terms of attitude toward schooling:

1. Single Sex – 97.3% (highest attitude)
2. Government – 94.3%
3. Denominational – 80.1 (lowest attitude)

\( H_4 \) There is no statistically significant relationship between gender, and performance at NCSE level 1

\( H_5 \) There is no statistically significant relationship between type of school, and performance at NCSE level 1

Spearman’s Rho Correlation was conducted to establish relationships between the variables of gender, and type of school, and performance in both examinations. The results show that there is a statistically significant relationship between the type of school (sig 0.000), and the gender (sig 0.000) of students who took the Language Arts examination at NCSE level 1.

\( H_6 \) There is no statistically significant relationship between gender, and performance at CSEC

\( H_7 \) There is no statistically significant relationship between type of school, and performance at CSEC
The findings show that there is a statistically significant relationship between the gender of students and type of school they attend (sig. 0.088), and gender and performance in the CSEC English A examination (sig. 0.000). There is also a statistically significant relationship between Gender of the student and type of school attended (sig. 0.090), and gender and performance (Sig. 0.000) in CSEC mathematics examination. As such, there is enough evidence to reject both null hypotheses.

### 6.5.1. Gender, NCSE and CSEC

In an attempt to further elaborate on the findings of the hypothesis testing, in being able to reject the null hypothesis, data was further analyzed to see exactly what the relationships were between the variables identified as significant. The data was standardized for accuracy in comparison of data due to the sample having more males than females. Results showed that there was an inverse relationship between male and female performance, where the more females attained passing grades, and peaked in the top 3 grade categories (A-C, or I-III), whereas more males attained failing grades and peaked in the lower grade categories (D-F, or III-ungraded).

### 6.5.2. Type of School, NCSE, and CSEC

The data showed that government schools are the lowest performing type of school, where there is an inverse relationship to the number of students that pass compared to the three other types of schools. It is interesting to note that the government single-sex school is one of the higher performing schools which is a sharp contrast to its sister government schools. The highest performing school in the single-sex denominational school. This finding was consistent throughout both subjects at both levels.

### 7.0. Discussion and Conclusion

These findings somewhat challenge the literature that takes a stance that low-stakes testing would be completely inverse to high-stakes testing, based on the fact that students are rarely motivated to succeed in low-stakes tests (Steedle, 2014). Inverse results would indicate that when students do not do well at NCSE, they go on to do well at CSEC. The study found that this inverse pattern exists in some results, particularly the higher grade levels in both Math and English/Language Arts. However, in the lower grade levels, the pattern of results becomes parallel indicating that the performance trends of students who do not do well at NCSE also do not do well at CSEC. These results are interesting to understand the attitude of students toward their schooling particularly in the
case of the individual standardized tests. The MOE identified that the rationale behind the use of the NCSE examination is a form of continuous assessment. With this rationale, there is evidence of much needed improvement in the administration and recording of data from the examination.

Teachers identified the challenge of consistently poor communication and education concerning the value and use of the NCSE examination. This has translated into much confusion and disinterest on the part of teachers and whole schools. Furthermore, the quality of the examination has been brought into question with teachers indicating not only spelling and grammatical errors, but even errors in the correct answers for the examination. Understanding that education does not occur in a vacuum, this study was also able to identify that the poor performance rate of students in the NCSE exam has contributing factors that include the confusion and lack of enthusiasm of their teachers, and their school administration.

The system of education in Trinidad and Tobago has not developed to the stage where results from low-stakes testing has a significant impact on the ranking and marketability of the schools, and as such, the implications of success rates at NCSE are not strong enough to propel an improvement in the quality of education at the school level. This has resulted in a poor participation rate of schools at NCSE, as well as poor performance on the part of the students. This study took a comprehensive approach to understanding student experience and its relationship to attitude toward schooling. The statistical limitations of the study restricted the researcher’s ability to answer the research questions as intended, however, gender and school type were found to be strong factors in a student’s attitude toward schooling.

This study found that there is a statistically significant relationship between gender, performance, and attitude toward schooling where female students tend to perform better than male students in English and maths at both NSCE and CSEC, and have a higher attitude rating. This is confirmation of existing research indicating that male students rarely compete on the same level as their female colleagues in academia. The female students’ performance average is not brought down by the results of the female students from government schools. This provides empirical evidence for the debate on the comparison of performance by gender in students in Trinidad and Tobago. Female students consistently outperform male students.
Based on the conceptual framework for this study, these findings could present a significant foundation for the evaluation of the psychological and physiological development of the student, and its impact on learning and assessment. According to theorists Bronfenbrenner (1989), Erikson (1968), and Piaget (1954), the stages of development both psychological and physiological must be evaluated before effective learning can take place. In this study, it can be determined that with such a consistent dichotomy between the male and female student in their results for the same standardised assessment, the assessment then does not factor the gender variable into its creation, design, and administration.

Findings consistently showed that the type of school had a strong statistical relationship to the performance of its students at both NCSE and CSEC. However, when measuring the student’s attitude toward schooling, it was found that the students in the lowest performing school still had a better attitude toward schooling. Examinee effort has been shown in the literature to be a major point of concern for testing of academic performance (Wise & DeMars, 2005). This study sought to attain some empirical insight into the attitude of students toward their schooling. Based on the findings of this study, there is enough evidence to suggest that in Trinidad and Tobago increased value needs to be placed on standardised tests, from administration through to completion.

Unfortunately, this study provides legitimacy for the findings of De Lisle (2012), who stated that reform strategies do not fit the demographic they intend to change. This is evidenced through poor communication of the roles and responsibilities to the key stakeholders; a lack of cultural ownership as evidenced through the errors and unreliability of the exams, as spoken to by some teachers in this study.

One of the greatest challenges that came out in this study, which eventually presented as a limitation to the study, was the lack of connectivity between the recording and administration of the two exams. The MOE has recorded the results of the two examinations on two separate databases with no identifying marks for the students who attempt the exams. This presents a major issue in the simple analysis of data, and production of reports on the performance of students throughout their secondary education in Trinidad and Tobago. Without this analysed data, the question is raised as to what evidence is used for education reform in Trinidad and Tobago. Test results cannot form any significant assessment at the student level for the development of continuous assessment and student
monitoring. De Lisle (2012) found that there is not always cohesion between the units in the MOE, and unfortunately, it appears that this is still the case.

Finally, there is still strong resistance and confusion among teachers as to what exactly the reform is doing for the country through education. Teacher perspectives in this study still mirror the feedback provided since De Lisle’s 2012 study which found that teachers do not buy into the reform strategies, and as such pass this resistance on to their students. Specifically cited by De Lisle (2012) were the many changes, lack of necessary training, and lack of effective communication. Based on the teacher perspectives presented in this study, these issues still exist. The issue of communication was particularly highlighted as teachers felt that information was not effectively disseminated as well as not being effectively received. The questions of value for the NCSE examination are still unanswered for stakeholders.

De Lisle (2012) identified the major problem in education reform on the whole, is where strategies were taken from outside sources and not effectively adapted to the local culture. This could be the source of the complete confusion that exists throughout the administration of the NCSE examination, which results in the ultimate devaluing of the exam by key stakeholders.

The value in NCSE has been continuously brought into question throughout the study by various stakeholders. The issues of value come from the following key areas as identified by the study: 1. Strategic intent of the examination; 2. Communication from MOE to stakeholders – particularly teachers; 3. Buy in from key stakeholders; 4. Performance of students in comparison to CSEC. 5. Transference of teacher dissatisfaction to students and parents.

Schrank (2016) indicated that one of the key ways to improve on the value of low-stakes assessment is to offer frequent testing as this is a more effective method of improving student learning and retention over time. The MOE can look into offering the NCSE assessment at various stages of the student’s secondary school experience, maintaining its low-stakes capacity while also providing a more comprehensive continuous assessment.

In establishing strategic intent, the MOE should look at the time of year during which the assessment is administered. Adjusting the timing of the assessment to be offered at the end of the second school term will allow for results to be published in time for students to use the results to
select their subjects for CSEC. Although not all subjects available at CSEC are covered by NCSE, career guidance can be sought based on the student’s experiences in the various subjects up until that point. This, again, provides for evidence-based practice in education, as well as value added to the assessment. This also creates a mapping of roles and responsibilities maximizing on the resources that the MOE places into the education system involving not only teachers, but also parents, social workers, and guidance counsellors in the process of NCSE, and therefore improving the quality of education.
8.0. References


Hislop, L. (2012). *The attitude of form three students, teachers, and parents to the administration of the NCSE level 1 examination in Trinidad and Tobago* (Unpublished Master’s thesis). University of the Southern Caribbean, Trinidad.


Pennebaker, J. W., Gosling, S. D., & Ferrell, J. D. (2013). Daily online testing in large classes:
Boosting college performance while reducing achievement gaps. *PLoS ONE*, 8(11), 1-6. doi: 10.1371/journal.pone.0079774


