THE UNIVERSITY OF THE SOUTH PACIFIC
LIBRARY

Author Statement of Accessibility

Name of Candidate : Gina Middleton

Degree : Master of Arts in Education

Department/School : School of Education

Institution/University : University of the South Pacific

Thesis Title : Assessment Conceptions and Practices among High School Teachers in Government-Funded Schools in Belize

Date of completion of requirements for award : August 2017

1. This thesis may be consulted in the Library without the author's permission. **Yes**/No
2. This thesis may be cited without the author's permission providing it is suitably acknowledged. **Yes**/No
3. This thesis may be photocopied in whole without the author's written permission. Yes/**No**
4. This thesis may be photocopied in proportion without the author's written permission.
   Part that may be copied:

Under 10%		40-60%	√
10-20%		60-80%	
20-40%		Over 80%	

5. I authorise the University to produce a microfilm or microfiche copy for retention and use in the Library according to rules 1-4 above (for security and preservation purposes mainly). **Yes**/No
6. I authorise the Library to retain a copy of this thesis in e-format for archival and preservation purposes. **Yes**/No
7. After a period of 5 years from the date of publication, the USP Library may issue the thesis in whole or in part, in photostat or microfilm or e-format or other copying medium, without first seeking the author's written permission. **Yes**/No
8. I authorise the University to make this thesis available on the Internet for access by authorised users. **Yes**/No

Signed: G Middleton

Date: 04 / Sept. /2017

Contact Address

7 Pikecresent Street, Buttonwood Bay
Belize City, Belize
Central America

Permanent Address

Same as Contact Address

ASSESSMENT CONCEPTIONS AND PRACTICES AMONG HIGH SCHOOL TEACHERS IN GOVERNMENT-FUNDED SCHOOLS IN BELIZE

by

Gina Middleton

A thesis submitted in fulfilment of the requirements for the degree of Master of Arts in Education

School of Education

Faculty of Arts and Law

University of the South Pacific

January 2017

DECLARATION

Statement by Author

I, Gina Anika Middleton, declare that this thesis is my own work and that, to the best of my knowledge, it contains no material previously published, or substantially overlapping with material submitted for the award of any degree at any institution, except where due acknowledgment is made in the text.

Signature: GMiddleton

Date: 24 August, 2017

Name: Gina Anika Middleton

Student ID No: S11119266

Statement by Supervisor

The research in this thesis was performed under my supervision and to my knowledge is the sole work of Gina Anika Middleton

Signature: ....LWaqailiti..........

Date: 24 August, 2017

Name: Dr. Ledua Waqailiti

Designation: Thesis Supervisor

DEDICATION

This thesis is dedicated to the memories of two beloved women; my great-grandmother, Matilda Stanford, and my grandmother; Constance Tasher. Their consistent love, attention and discipline during my childhood had moulded me into the woman I am today.

ACKNOWLEDGEMENT

Foremost, I would like to acknowledge the continuing presence of God in my life from the start to the end of this journey. He had kept me and revealed things to me during this journey that did not only broaden my knowledge on the research issue, but sharpened my character.

I thank my supervisor, Dr. Ledua Waqailiti, for her continuous support throughout the course of my study. She supported and motivated me to complete this thesis on time and was always there to guide me through the process. Thanks also goes out to my co-supervisor, Mr. Hem Dayal, who provided insightful questions and thoughts on my topic and pointed me to relevant literature and thinkers in the field. I also thank the School of Education's secretary; Ms. Ulina Tuilau, and the Head of School's administrative assistant; Mrs. Vilma Ratumudu. They were always there to answer my administrative questions, and when they could not, they directed me to the right people. Heartfelt thanks to Mrs. Vilma Ratumudu who always greeted me with a bright smile and willingly helped out with whatever concerns I had. I also acknowledge the university research statistician; Mr. Karuna Reddy, who assisted me in the analysis of my data.

I was supported in full by the Caribbean-Pacific Island Mobility Scheme (CARPIMS) and I take this opportunity to thank the European Union for the opportunity provided to me to complete this degree at the University of South Pacific.

Special thanks goes to my friends and fellow colleagues. Friends are indeed special gifts from God and he had blessed me with an amazing group of friends during this journey. Thanks goes out to Shonell, Arieni, Cleopatra, Geon, Mia, Catherine and Deshawn for their support and friendship during this time. Special acknowledgement goes to Mrs. Carol Lord who willingly read over and provided critical comments on my chapters. Special acknowledge goes to all the teachers who took part in the study.

Last, but certainly not least, I would like to acknowledge my family. To my mother; Jennifer Felix, my cousin; Karen Tasher, and my significant other; Darrell Reyes, for

ensuring that my son did not lack the attention and love due to him in my two years of absence from his life. Thanks also for their unconditional love towards me and their firm belief in my ability to succeed.

ABSTRACT

This study examined assessment conceptions and practices among high school teachers in four core subject areas; Mathematics, English Language, Social Studies and Science. Specifically, the study sought to determine the types of assessment practices used by teachers, the frequency with which they were used, how teachers used the information gathered from those assessments and their assessment conceptions. The study made comparisons between teachers' assessment conceptions and practices and teachers' characteristics such as gender, subject taught, teaching experience and assessment training. Examination was also done on the relationship between teachers' assessment conceptions and their assessment practices. The focus was on in-service teachers in six government-funded high schools in the Belize district who completed a questionnaire on their assessment conceptions and practices. Analysis from the data revealed that teachers used a variety of assessment practices but tended to favour summative / traditional assessment practices. Their assessment practices were in line with their assessment conceptions since teachers simultaneously agreed with improvement and student accountability conceptions. Significant differences were yielded between assessment conceptions and practices and selected teachers' characteristics. Significant results were also produced between assessment conceptions and assessment practices indicating that teachers' beliefs about assessments are in line with their assessment practices.

ABBREVIATIONS

AfL	–	Assessment for Learning
AoL	–	Assessment of Learning
ATLIB	–	Association of Tertiary Level Institutions of Belize
BJAT	–	Belize Junior Achievement Test
CSEC	–	Caribbean Secondary Education Certificate
CXC	–	Caribbean Examination Council
FA	–	Formative Assessment
MOE	–	Ministry of Education
PSE	–	Primary School Examination
SBA	–	School Base Assessment

GLOSSARY

1. Assessment: The process of gathering evidence and interpreting that evidence in the light of some defined criterion in order to form judgement.
2. Formative Assessment: In this study, formative assessment and assessment for learning were used interchangeably and both referred to activities undertaken by teachers and students that provide information to be used as feedback to improve and modify learning and teaching.
3. Summative Assessment: Used to measure what students learn at the end of a unit, to promote students from one level to the next, to meet required standards for completion and certification, to enter certain occupations or professions and to enter further educational institutions.
4. Assessment Practices: Manners in which teachers conduct assessments such as observations and tests.
5. Classroom assessment: The planned method or strategy used in a classroom to establish the level of student's learning.
6. Conceptions of Assessment: In this study the word conception meant the same as the word belief and it reflected teachers' thinking about the purpose of student assessment. It is also used as a frame of reference to interpret assessment policies and regulations (Brown, 2006).
7. Government Funded Schools: Schools that are fully funded by the government of Belize. These schools are managed by members of a board appointed by the Ministry of Education.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

1.0 OVERVIEW OF THE STUDY

1.1 INTRODUCTION

Forty six years ago, Paulo Freire in his book, “Pedagogy of the Oppressed”, compared education to a banking process in which teachers were the depositors and students were the depositories (Freire, 1970). He showed teachers passing on content to students who passively received it just as containers received water. Twenty eight years later, Black and Wiliam (1998) drew attention to the happenings inside the classroom with the black box analogy. The analogy showed that what happened inside the classroom was unknown to many; yet, it significantly impacted students’ learning. As Black and Wiliam studied what was inside the black box, they saw classroom assessment as the tool that lays at the heart of effective learning.

Many studies and reviews have since illustrated the significant impact that classroom assessment has on teaching and learning (Black & Wiliam, 1998; Brookhart, 2007; Crossouard & Pryor, 2012; Heritage, 2007; Organisation for Economic Co-operation and Development (OECD), 2005; Wiliam, 2013; Wiliam, Lee, Harrison & Black, 2010). Several assessment studies have shown that purposeful assessment enhances students’ learning and achievement; yet, teachers are still assessing learning the way their predecessors did sixty years ago (Clark, 2012; Heritage, 2007; Stiggins, 2005; Volante, Drake & Beckett, 2010). Educators seem to be obsessed with traditional assessment practices such as tests, quizzes and examinations which remain at the forefront when it comes to classroom assessment (Heritage, 2007; Stiggins, 2002). This culture of testing dominates teaching and assessment with teachers giving high priority to the collection of marks rather than the analysis of students’ work to assess whether learning has occurred (Black & Wiliam, 1998).

The contrast between what research recommends and what is actually practiced is what propels researchers to call for more studies on teachers’ conceptions (beliefs) of assessment (Azis, 2012; Brown, Kennedy, Fok, Chan, Yu, 2009; Segers & Tilleman, 2011). It has continuously been suggested that teachers’ conceptions of assessment may be related to their assessment practices. Nespor (1987) was among the first to claim that "to understanding teaching from teachers' perspectives we have to understand the beliefs with which they define their work" (p. 324). This study; therefore, not only examines the quantitative findings of secondary teachers' assessment conceptions and practices in Belize, but also looks at the relationship between conceptions and practices. This first chapter presents the background of the study, specifies the problem of the study and describes the significance of the study. The chapter concludes by noting the limitations of the study and defining some significant terms.

1.2 BACKGROUND OF THE STUDY

Traditional model of teaching and testing had dominated education in the twentieth century (Berry & Adamson, 2011; Gipps, 1999; Shepard, 2000). This model, based on behaviouristic and measurement theories, created a teaching / learning environment that placed teachers in the centre of the equation (Gipps, 1999; Shepard, 2013). Teachers were seen as transmitters of knowledge of which students passively received. Knowledge was seen as a distinct body of information that was to be broken down into pieces and passed on to students. Assessment sat on the outside and was aimed at measuring the amount of knowledge students' acquired. Contrary to this traditional model, there is an emerging model that asserts that assessment should be done in collaboration with teaching and should be concerned with improving students' learning rather than merely measuring it (Berry & Adamson, 2011). This is in line with cognition, constructivism and sociocultural theories which set forth that individuals construct their own knowledge by doing their own interpretation and organization within a specific social context using cultural tools (Berry & Adamson, 2011; Krause, Bochner, Duchesne & McMaugh, 2010). Despite years of appeal and a huge body of research on the topic, the idea that assessment should work in line with teaching to improve students' learning is still not full endorsed by teachers (Heritage, 2007).

Research has shown that teachers use a variety of assessment practices, but traditional assessments are used more frequently (Azis, 2015; Bol, Stephenson, O'Connell & Nunnery, 1998; Frey & Schmitt, 2010; Ogunkola & Clifford, 2013; Stiggins & Bridgeford, 1985; Suah & Ong, 2012). Assessments that are used formatively are considered the best approach to evaluate and enhance students' learning. These require teachers' observation and professional judgement to make decision regarding students' learning (Cizek, 2010; Popham, 2005). They emphasize knowledge construction and problem solving and encompasses several different practices including performance assessment, observation techniques, peer and self-assessments, feedback, questioning and portfolios (Bol et al., 1998; Cizek, 2010; Zhang & Burry-Stock, 2003). These formative assessments are seen as the best approach because they promote higher standards of teaching and more powerful learning. It goes beyond traditional assessment by motivating students to take more responsibilities for their learning, making assessment an integral part of learning and creating active learners rather than passive learners (Cizek, 2010; Hargreaves, Earl & Schmidt, 2002; Shepard, 2013). Even though formative assessments are widely acclaimed as being the best practices for improving student's learning, studies have shown a low frequency of its usage and understanding among teachers (Clark, 2012; Ndwandwe and Dlamini, 2014; OECD, 2005; Song & Koh, 2010; Stiggins & Bridgeford, 1985; Volante, Drake & Beckett, 2010). In 1985, Stiggins and Bridgeford revealed that even though three-quarters of teachers used some form of performance based assessment, only half of them felt comfortable using it. It was further noted that such forms of assessment were used less as grade level increased. In 2014, Ndwandwe and Dlamini discovered that even though teachers highly endorsed alternative methods of assessment, traditional methods were used when they were being observed.

Traditional assessments refer to the standardized classroom assessments that feature close-ended items, multiple choices, true and false items, checklists and fill in the blanks (Bol et al., 1998; Looney, 2011; Volante & Beckett, 2011; Zhang & Burry-Stock, 2003). These assessments only address distinct measurable behaviours and focus on the end result of learning rather than the process (Looney, 2011; Volante & Beckett, 2011). In 1985, Gullickson discovered that traditional paper and pencil tests were the most common methods used by teachers. Twenty five years later, research done by Frey and Schmitt (2010) also yielded similar results.

Contrary to findings that revealed teachers frequently using traditional assessment practices, studies on teachers' assessment conceptions suggest that they highly value assessment practices that aim at improving teaching and learning. Teachers have been found to agree highly with improvement conceptions of assessment (Azis, 2015; Brown, 2004; Brown et al., 2009; Brown, Hui, Flora & Kennedy, 2011a; Brown , Lake & Matters, 2011b; Brown & Michaelides, 2011; Segers & Tillema, 2011) suggesting that they believe greatly in assessment practices that allow them to improve their teaching and increase students' learning. The problem is that studies also found teachers agreeing with practices that make students' accountable for their learning (Azis, 2015; Brown et al., 2009; Brown & Michaelides, 2011; DeLuca & Lam, 2014; Ndalichako, 2015). These inconsistencies among findings have been attributed to a country's societal and educational contexts (Barnes, Fives & Dacey, 2015; Brown et al., 2009; Brown & Hattie, 2009; Black & Wiliam, 2009). For instance, a stronger correlation was found between improvement and student accountability conceptions in Hong Kong compared to New Zealand and Queensland suggesting that teachers in Hong Kong highly believe that improvement in teaching and learning occurs best when students are held accountable for their learning (Brown & Hattie, 2009). Such differences in beliefs among the teachers were attributed to the cultural and educational contexts in their countries since there had not been a history of high-stake testing in New Zealand and Queensland as compared to Hong Kong (Brown & Hattie, 2009).

Few studies concurrently measure teachers' assessment conceptions and practices to gain better understanding of how the two relate to each other (Barnes, Fives & Dacey, 2015; Ndalichako, 2015; Segers & Tillema, 2011). The few existing studies show inconsistences on the matter. Some studies found that teachers' conceptions are in fact in line with their practices (Adam & Hsu, 1998; Brown et al., 2009; Calveric, 2010), others found the opposite (Azis, 2015; Chew & Lee, 2013; James & Pedder, 2006). Examining teachers' conceptions of assessment may shed light on the assessment practices they utilize.

In Belize, there are little known empirical evidences available on teachers' assessment practices and basically none on their conceptions of assessment. The little that is known is mostly based on unreliable and unscientific evidence (Wolff & Castro, 2000). Reports on Belize's education system have indicated that teachers' practices are mainly based on outdated pedagogical models and are usually teacher-centred (Naslund-Hadley, Alonzo, Martin, 2013). One study that examined whether social constructivist activities could improve mathematical competency of grade five students in Belize found that students from the controlled group demonstrated higher gains than students from the experimental group (Brown-Lopez, 2010). One of the reasons provided for such finding was that teachers lacked the proper training in using constructivist-based activities. The study also noted that the teachers' practices were reflections of their teaching and learning experiences and not of innovative teaching activities.

1.3 STATEMENT OF THE PROBLEM

It has been estimated that teachers spend a huge percentage of their time on assessment related activities (Robinson, Maldanado, Whaley, 2014; Stiggins, 2014). This has caused educators to look upon classroom assessment as an instrument to meet the goals of high performance, high equity of students' outcomes and lifelong learning skills (OECD, 2005). Classroom assessment provides information to improve teaching and learning which in turn empowers and motivates students (Alkharusi, 2010; Nitko, 2004). Furthermore, when assessment is improved, it has been shown to help low achievers and students with disabilities to improve their learning (Black & Wiliam, 1998; Crossouard & Pryor, 2012; James & Pedder, 2006; OECD, 2005). Nonetheless, even though studies have shown the far-reaching impact of assessment practices, it is rarely used effectively in the classroom due to different barriers (Black & Wiliam, 1998; Bol et al., 1998; Brookhart, 2007; Clark, 2012; Duncan & Noonan, 2007; Stiggins, 2014).

One of the main barriers to effective use of assessment is teachers' conceptions of it. Teacher's conceptions and values about teaching underpin and influence their practices in the classroom (Cizek, Fitzgerald & Rachor, 1995; Scott, Webber, Aitken

& Lupart, 2011) and thus prevent them from using assessment effectively in the classroom (Black & Wiliam, 1998; Stiggins, 2014). Studies have also shown conflicting views between what teachers believe to be of value for student's learning and what they actually practice in the classroom (Chew & Lee, 2013; Ndwandwe & Dlamini, 2013). Others studies have shown how teacher's psychological beliefs influence assessment practices. These studies make recommendations for further studies on teachers' conceptions of assessment (Alkharusi, 2010; Brown et al., 2009; Erkmen, 2012; Song & Koh, 2010; Young & Kim, 2010; Zhang & Burry-Stock, 2003).

The underlying effect of a country's social and educational contexts, the inconsistencies among findings, the scarcity of empirical evidences in Belize and the few existing studies that simultaneously examine the assessment conceptions and practices all make the issue worthy of study. The information gathered will greatly add to the discourse and can positively impact the practice and the teaching learning process.

1.4 PURPOSE OF THE STUDY

This research sought to examine teachers' assessment conceptions and practices in Belize. It further attempted to determine the types of assessment strategies used by teachers, the frequency with which they were used, how teachers used the information gathered from those assessments and their conceptions about assessment. The study made comparisons between teachers' assessment practices and conceptions and their characteristics such as training, subject taught, years of experience and gender. The study also examined the relationship between teachers' conceptions of assessment and their assessment practices.

1.4.1 Research Questions

To fulfil the aims of the research, the following research questions were used to guide the process:

1. What assessment practices are used by teachers in government funded high schools in Belize?
2. What are teachers' conceptions of assessment in government funded high schools in Belize?
3. What are the differences between assessment conceptions and practices and teacher characteristics such as gender, training, teaching experience and subject taught?
4. How do teachers' assessment conceptions relate to their assessment practices?

1.4.2 Overview of the Methodology:

The study sought to find out assessment practices and conceptions among high school teachers. A quantitative methodology through the use of a survey design was used in the research. A quantitative methodology allowed the researcher to not only study the overall results, but to examine the variation of results based on teachers' characteristics. The survey design was ideal because it allowed for the description of people's behaviour, perceptions, opinions and beliefs about current issues in education (Lodico, Spaulding & Voegtle, 2006). A cross-sectional questionnaire was administered to teachers currently teaching in government funded high schools in the Belize district. The target population was teachers teaching in government funded high schools in the Belize district. All teachers who taught mathematics, English Language, social studies and science were selected to participate in the study.

The questionnaire administered to teachers consisted of two parts; assessment conceptions and assessment practices. Items for assessment practices were gathered from the literature and were reviewed by experts in assessment, while the items for assessment conceptions came from existing scale created by Brown (2006). Analysis of the data was done using the Statistical Package for the Social Sciences (SPSS).

Descriptive statistics in the form of means, standard deviation, frequency and percentage were used to summarize the trend and tendencies of teacher's assessment practices and their conceptions of assessment, while t-test and correlation coefficient were used to explain the relationship between teachers' conceptions and their assessment practices and between practices and conception and their characteristics.

1.5 SIGNIFICANCE OF THE STUDY:

Raising standards in education is a worldwide issue (Black & Wiliam, 1998; Stiggins, 2014; Wolff & Castro, 2000). Focusing on standardized testing and change management will do little to enhance students' achievement, the focus should be placed on the happenings inside the classrooms (Black & Wiliam, 1998; Stiggins, 2014). Classroom assessment is looked upon as a major means to improve teaching and learning (Black & Wiliam, 1998; McMillan, 2013; Stiggins, 2014). Through the implementation of thorough and regular assessment practices, teachers can improve students' learning and achievement (Black and William, 1998; McMillan, 2013). The right and proper usage of assessment practices can impact learning not only for high achieving students, but also for low achieving and disabled students (Black & Wiliam, 1998; Crossouard & Pryor, 2012, Wiliam et al., 2010).

Besides improving learning, the use of effective assessment prepares students for the 21st Century workforce (Chew & Lee, 2013; Stiggins, 2014; Wolff & Castro, 2000). The purpose of school has evolved and schools are now requireD to equip students with prized skills such as problem solving, collaboration and creativity that will enable them to survive in the 21st century (Chew & Lee, 2013; Stiggins, 2014). Teachers' assessment practices need to accommodate these changes. Many have advocated for the use of assessment practices that are grounded in the constructivist approaches. However, such practices are used at a lesser rate by teachers who seem to favour practices grounded in traditional approaches. This detachment between literature and practices could be understood by examining teachers' conceptions of assessment (Alkharusi, 2010; Brown et al., 2009; Erkmen, 2012; Song & Koh, 2010; Young & Kim, 2010; Zhang & Burry-Stock, 2003)

It is important to study teachers' conception of assessment because teachers' psychological beliefs have been shown to influence the use of assessment practices (Alkharusi, 2010; Song & Koh, 2010; Young & Kim, 2010; Zhang & Burry-Stock, 2003). Evidence has shown that what teachers teach and what students learn are heavily influenced by teachers' conceptions (Brown, 2004). Results from studies done on teachers' assessment practices are usually ignored because such studies fail to provide reasons as to why teachers used such practices (Brown et al., 2009). For instance, Borko et al. (1997) discovered that new ideas and practices were ignored by teachers due to incompatibility with their existing beliefs (as cited in Young & Kim, 2010). Thus, the study of teachers' assessment conceptions will help to explain why changes in assessment practices are implemented differently at the national and school levels by teachers (Segers & Tillema, 2011). There are few research studies that examine teachers' conception about assessment (Song & Koh, 2010; Zhang & Burry-Stock, 2003), much less research that relate teachers' conceptions with their assessment practices (Brown et al., 2009; Calveric, 2010).

With the outstretch of globalization and the changing characteristics of the world's economies, developing countries will require a labour force equipped with not only mathematics and communication skills, but with the ability to work together and think creatively. Many countries in Latin America and the Caribbean have been working to enhance the quality of secondary education which is considered inadequate when compared to international standards. Their inadequacy is due to outdated curriculum and poor pedagogical training of teachers (Wolff & Castro, 2000). Majority of the evidence on teachers' classroom practices are based on anecdotal information rather than empirical evidences (Wolff & Castro, 2000). As Belize pushes reforms to improve curriculum and assessment at the secondary level, empirical evidences will be needed to support those initiatives. This research will provide administrators and policy makers with empirical evidences on teachers' assessment conceptions and practices so that proper strategies and programmes can be implemented at the national and school levels.

The results from this study can also be used to improve teachers' training programmes in the country. A high agreement with using assessments to keep student accountable and a low agreement with using it to improvement teaching and learning might be indicating that modification is needed in teachers' training programmes. The examination of teachers' assessment conceptions and practices in relation to their gender, subject taught, assessment training and teaching experience provides beneficial information that could aid in the improvement of teacher training. For instance, it has been noted that assessment practices vary across subject areas (McMillan, 2013) suggesting that this fact should be kept in mind when creating and structuring assessment courses. The examination of the relationship between measurement course and teachers' assessment conceptions and practices provide interesting findings on the impact of measurement courses taken by teachers in their training programmes. For instance, Armstrong (2011) discovered that primary school teachers in Belize with associate degrees in primary education held the highest level of assessment understanding and the highest usage of positive assessments when compared with teachers who had bachelor degrees in primary education. The difference was attributed to the quality of teaching since it is believed that teachers with an associate degrees were exposed to more qualified lecturers in their teacher education programmes than those in bachelor degree programmes.

1.6 LIMITATIONS:

Certain aspects of the study need to be taken into consideration before full acceptance of the findings is taken.

1. The target population of the study was secondary teachers in government-funded high schools in the Belize district who were teaching either of the four core subject areas; mathematics, English Language, social studies or science. One hundred and ninety one teachers made up the entire population while eighty nine teachers made up the target population. It would have been ideal if the entire eighty nine teachers completed the questionnaire since this would have created a greater statistical power when analysing the data. However, this was not the case since only fifty eight out of the eighty nine teachers completed the questionnaire even though the entire population was invited.

A possible reason for this might have been due to the use of a research assistance for distribution and collection since the researcher was far away from the study site. The researcher might have been able to persuade more teachers to participate in the study had she been present. Nonetheless, the absence of the researcher and the fact that the data was analysed some distant away from the study site may have contributed to more sincere and honest responses.

2. Due to the small sample population, the issue of type one and type two errors may be at play in regards to findings based on teachers' characteristics. The sampling sizes for the different group of teachers' characteristics varied and in some cases were too small. The discovery of either a statistical significant result or the lack of a statistical results based on teachers' characteristics could be attributed to the small sampling sizes.

3. The time period of distribution and collection may have affected the amount of participants and their responses. The survey was distributed and collected within a prescribed timeframe. This is mainly due to the differences in the school year between Fiji and Belize. While the school year starts in February and ends in November in Fiji, such is not the case in Belize where the school year starts in September and ends in June. That meant that distribution and collection had to occur at the end of the school year which was the most stressful time for the teachers due to internal and external examinations. Thus, some of the respondents might have felt overburdened or stressed while completing the questionnaire.

4. Another aspect that might have impacted the findings is the use of a questionnaire. Questionnaire was the ideal method in this case since it is economical due to the low cost and it can be administered quickly. However, results are based on self-reported data which might have been influenced by social desirability. Respondents might have chosen responses that they thought were acceptable or right. The use of a questionnaire also limited the total understanding of teachers' actual assessment practices. Nevertheless, the fact that each respondent remained nameless and knowing that data would

be seen mainly by individuals in a different country might have prompted the majority of the respondents to be as truthful as possible.

1.7 ORGANISATION OF THE STUDY

This chapter provides an overview of the study. The next chapter, Chapter Two, gives a description of the background of Belize and its education system. It also highlights current policies regarding secondary education in Belize.

Chapter Three provides a discussion and analysis of the relevant literature pertaining to teachers' assessment practices and their conception of assessment. It starts with an explanation of the theoretical framework encompassing the study which is followed by sections that frame the issues of the study. These include an overview of the historical context and tension within the field of assessment, defining classroom assessment and assessment conceptions. The subsequent sections provide detailed analysis of the major themes and empirical evidences regarding assessment conceptions and practices. This is followed by an analysis and discussion of the relationship between conceptions and practices. The last section deals with teachers' assessment practices and conception in Belize.

Chapter Four details the research design developed and used in the study. This research was concerned with studying teachers' assessment practices and beliefs from a quantitative methodology through the use of a survey. This chapter describes the research design, population, sampling, instrument, data collection and data analysis.

Chapter Five gives an analysis of the results. It presents an analysis of the data by research questions and also provides a summary of the results.

Chapter Six presents an analysis and discussion of the main findings of the research and makes references to relevant literature.

Chapter Seven draws conclusions from the findings, provides some implications and suggestions related to the findings and presents some directions for future research studies.

CHAPTER TWO

2.0 BACKGROUND OF BELIZE AND THE EDUCATION SYSTEM

2.1 INTRODUCTION

Education is seen as the key component required to cultivate the human resources needed for economic development in Belize. The importance placed on education is seen in the fact that one of the two pillars that Belize's national development plan, Horizon 2030, is built upon is education (Barnett, Catzim-Sanches & Humes, 2010). This section serves to provide a brief explanation of Belize's geographical, economical, political and societal context. It will also provide an insight into the education system of Belize focusing specifically on secondary education and assessment.

2.2 CONTEXT OF BELIZE

2.2.1 Location and Geographical Features

Belize, formerly known as British Honduras, is located in Central America (see Figure 2.1) but shares ties with other English-speaking Caribbean countries. It is located on the coast of the Caribbean Sea and is neighbour to Mexico in the north and Guatemala in the south and west. Once a colony of Great Britain, Belize is the only English-speaking country within Central America. It has a territory of 22,960 square kilometres that includes islands, reefs, coastline, rivers, lakes, caves, swamps, mountains, hills and forests. Because of its extensive coastland areas, Belize is susceptible to natural disasters such as hurricanes, tropical storms and flooding. These factors are impediments to the development of the country.

Figure 2.1. *Map Showing the Location of Belize in Central America*

The country is divided into six districts; Corozal, Orange Walk, Belize, Cayo, Stann Creek and Toledo. Corozal and Orange Walk are located in the northern part of the country and Stann Creek and Toledo are located in the southern part. The capital, Belmopan, is located in the Cayo district in the western part of the country. The Belize district, which is located along the eastern coastline, includes Belize City, the most urbanized and populated area in the country.

2.2.2 Population

The population of Belize is little over 350,000 people of mixed ethnicities and cultures. It consists of the Creoles (descendants of the British settlers and / or African slaves), Mestizos (mixture of Spanish and Mayan), Garifuna (descendants of the Carib Indians), Mayans (the indigenous people), Mennonites, Lebanese, East Indians, Chinese and others (Statistical Institute of Belize (SIB, 2013). The country was once predominantly creole, but due to the high rate of immigration from neighbouring countries such as Guatemala, Honduras, Mexico and San Salvador, the dominant race as noted in the 2010 Census is Mestizo (SIB, 2013). Besides having a diverse population, Belize has an ageing population with a dependency ration of 66% and it is projected to increase in the future (Caribbean Development Bank (CDB), 2010). Furthermore, the poverty and indigent rates in Belize have significantly risen from 2002 to where there is currently 30% - 40% of the population of Belize living in poverty (CDB, 2010; SIB, 2013).

The diversity in ethnicities and cultures, the aging of its population and the high poverty rate call for a sound education system. This system should emphasize diverse teaching practices to cater to its diverse student population. It will also need to produce skilled and knowledgeable workers to develop the economy due to its future dependency on the working class population. Lastly, the high poverty rate could be related to education in the fact that half of Belize's households where the head did not complete primary school are poor, compared to only 15% of households where the head has some secondary or tertiary education (CDB, 2010). This makes education even more important in producing a skilled working force that will be able to access employment opportunities and encourage investment to build economic growth in order to reduce Belize's poverty rate (CDB, 2010).

2.2.3 Economic System

Belize's economy is primarily based on agriculture, fishery and tourism (Carneiro, 2016). Agriculture was once the backbone of the country's economy as it produced the highest among of wealth, employed many people and fed the country; however, the tourism industry has taken the lead as it provides for 28% of employment in Belize and contributes to 21% of GDP (Carneiro, 2016). The main products that are exported are sugar, citrus and banana. Fishery is another vital sector of the economy in that it provides 3.1% of the Gross Domestic Product (GDP) (SIB, 2013). With discovery of oil in 2005 and the boost in the tourism industry, Belize's economy has transformed in the last two decades. Nonetheless, it is extremely vulnerable to the external environment and natural disaster.

Since Belize imports more than it exports, it is not only dependent on products from the world market for local consumptions, its primary industries are also depended on the world market (Carneiro, 2016). This means that a crash on the world market, may also mean a crash on Belize's economy. In addition, its exposure to natural disaster also makes the country's economy vulnerable. For instance, the economy went down from 4.1% in 2014 to 1.7% in 2015 due to weather-related incidences that caused poor agricultural outputs (World Bank, 2015).

Over the past two decades, Belize has consistently allocated a high percentage of its Gross Domestic Product (GDP) to education. Despite that, the rewards in terms of quality, access and equity are quite disappointing (Naslund-Hadley et al., 2013). This highlights the need for a system that emphasizes quality so that it can positively impact the economic development of the country.

2.2.4 Political System

Upon gaining self-government in 1964, the British Westminster System was adapted and continues to be used after Belize gained its independence in 1981. The country chose to remain a member of the British Commonwealth mainly to ensure protection by British forces from Guatemala, its neighbour to the west and south, who has an ongoing claim against the country.

The Westminster System refers to a democratic system of government in which the British queen or king, represented by the Governor General, is the head of the country. A bicameral system makes up the legislative branch of government which consists of an elected House of Representative and an appointed senate. Both houses make up what is known as the National Assembly. The Prime Minister, who is the head of government must command the support of the majority of members in the House of Representatives. Bills and Laws are tabled and debated in the House of Representatives and then ratified and confirmed in the Senate. The House of Representatives consists of thirty one elected representatives and the Senate consists of twelve members that are appointed by the Governor General on the advice of the Prime Minister, the leader of the opposition and other entities.

Two main political parties dominate the political landscape in Belize; the People United Party (PUP) and the United Development Party (UDP). The PUP won all general elections until 1984 when they were defeated by the UDP for the first time in thirty years. Since then the two parties have exchanged power over the years at each election until 2003 when PUP was re-elected for a second term. The current government, UDP, has been in power since 2008. Politics plays a major role in Belize since many of the electorates are affiliated with either of the two parties and usually expect certain benefits once their party takes control of the government (CDB, 2010).

2.2.5 Crime Problem

The crime rate in Belize has increased from 16% per 100, 000 inhabitants in 2002 to 41% in 2010 (SIB, 2013). In 2012, the murder rate was estimated to be 42.5% per 100,000 inhabitants with Belize City recording the highest rate estimated to be 106 per 100,000 inhabitants (Carneiro, 2016). A study conducted by Gayle and Mortis (2010) reported that 99% of Belizean youths are exposed to violence. This study also pointed out that Belize has the 6th highest crime rate in the world.

Dowd and Aiello (2009) noted that educated people are less likely to engage in criminal activities (as cited in Carneiro, 2016). This means that quality education should play a significant role in reducing Belize's crime rate. In fact, the improvement of schools, especially those located on the Southside of Belize City, is one of the main strategies recommended by several reports to reduce the crime rate in Belize (Barnett et al., 2010; Carneiro, 2016; CDB, 2010). HGL (2010), for instance, recommended that the profile of schools be risen, relationships between teachers and parents be developed and outreach programmes between the communities and schools be encouraged in order to tackle the issue of crime in Belize City.

The next part of this section looks at the education system in Belize. It will look especially at secondary education and the policies that are in place with regards to assessment practices.

2.3 THE EDUCATION SYSTEM IN BELIZE

2.3.1 *Brief Background*

During the period of colonization, education was placed in the hands of missionaries who arrived in 1828 and the Jesuits who arrived in 1850. These religious entities continued to dominate the policy making and management of education in Belize for many years. However, political and social changes during the 1960s – 1980s gave rise to a nationalist movement that called for the restructuring of education with the aim of making it more "Belizeanized" (Thompson, 2000).

2.3.2 Distribution and Classification of Schools

Belize has a total of 607 schools; 222 preschools, 300 primary schools and 56 high schools (Ministry of Education (MOE), 2015a). Eighty percent of schools are privately owned in Belize but are subsidized by the government. Majority of schools are located in the districts of Belize (142) and Cayo (139). Of the fifty six high schools in Belize, only seventeen are owned by the government.

The education system of Belize is divided into four levels; preschool, primary, secondary and tertiary. Children aged 3 – 5 usually attend preschool. From age 6 – 13 children attend primary school and partake in two standardized examinations. The Belize Junior Achievement Test (BJAT) is taken in standard three (which is equivalent to grade five in the U.S. education system) and it is used to measure the students' performance at that level (Thompson, 2000). The second standardized test is the Primary School Examination (PSE) which is taken in standard six (equivalent to grade eight in the U.S.). The two main purposes associated with PSE are to provide certification of students at the completion of primary school in four content areas (Mathematics, Science, Social Studies and English Language) of the primary curriculum and to form the basis for educational decision- making at the national, school and classroom levels (MOE, 2015b). PSE is not used for graduation purposes nor for selection to secondary school, graduation and selection are mainly done based on students' internal summative grades. Thus, both external assessments taken at the primary level are low-stake exams.

After primary school, students transition into high schools where teenagers mostly of ages 14 – 17 are enrolled. Secondary education is divided into four forms that are completed within four years. Two standardized examinations are taken at the end of secondary school.

2.4 SECONDARY EDUCATION IN BELIZE

2.4.1 *Brief Background*

Majority of the early reforms in education during the period of 1960s – 1990s were concentrated at the primary level with little consideration given to secondary education. In fact, secondary education was not introduced in Belize until the last two decades of the nineteenth century (Bennett, 2008). For a long time secondary schools followed a British or an American curriculum. Traditional assessments such as drill and practice were often use with the aim to make students remember as much facts as possible in order to pass external examinations (Bennett, 2008). Those who passed those examinations were offered scholarships and entrance into external universities (Thompson, 2000). However, with the establishment of the Caribbean Examination Council (CXC) in 1972, students started to sit the regional examinations offered by CXC which consequently saw secondary schools adapting the Council's syllabus (Bennett, 2008).

2.4.2 *Assessment at the Secondary Level*

By the time the new millennium came around, countless curriculum reforms occurred at primary level, but nothing much was been done with regards to assessment and evaluation at the secondary level. Schools continue to develop their own internal policies and practices for assessing and evaluating students' achievement as specified in Section 126 (1) of the Education and Training Act 2010. This means that schools may differ in their practices and reporting of assessment. Furthermore, even though a secondary school curriculum was created in the early nineties, this curriculum was (and still is) for lower secondary (form 1 and 2) and few teachers are aware of its existence, much less follow, that curriculum. Thus, CXC's syllabi continue to be the main curriculum utilized in secondary schools across Belize and the Caribbean Secondary Examination Council (CSEC) examinations, along with the Association of Tertiary Level Institution of Belize (ATLIB) examination, are the only two standardized examinations taken at the end of secondary school.

CSEC offers two schemes to students; general and technical proficiencies, across various subject areas. Many of the examinations offer a wide range of measurement techniques such as final written examination and internal school based assessment (SBA) (Gropello, 2003). However, CSEC is not compulsory; thus, not all secondary students take the examinations. Furthermore, despite the fact that students' performance are based on a six-point grading scheme on overall performance (points range from I - comprehensive grasped to VI - limited grasp) and a subject profile (points range from A – outstanding to F – weak), no real initiative has been taken to evaluate grades received by students (Gropello, 2003). In the same way, the ATLIB examination which consists of two papers; Math and English, is used to determine placement in tertiary institutions and is usually not taken by students who have no desire to continue their studies (Bennett, 2008). These reasons make it difficult to know the exact performance of individual schools and secondary students' performance based on these examinations.

2.4.3 Secondary Teacher Training

Of the 496 secondary teachers in Belize district, only 35.1% of them are trained (MOE, 2015a). Before 1954, no institution offered formal teachers' training in Belize; however, there are currently several institutions in Belize that offer teachers' training programme. The rise in institutions offering training programme has to do with demand for trained teachers since the MOE recently legislated that all in-service teachers must be trained.

The main institution offering teachers' training programme is the University of Belize (UB) which offers two routes for secondary teachers to gain their certification in education. The first route is a four semesters Bachelor in Secondary Education with concentration areas being Math, English, Biology and Business. During this programme, students will take eight courses, including one course in classroom assessment, introducing teachers to the main issues of pedagogy. Teachers also need to complete one semester of internship. For the second route, teachers who already possess a bachelor degree in their specialized areas, proceed to take the diploma in education which offers a two semester programme and follows the same component of internship as mentioned before.

2.4.4 *Education Reform and Policies*

As mentioned, Belize has consistently invested a huge portion of its GDP into education believing that the key to development is through human resources. Despite this, Belize is not adequately educating its people and the payoff in terms of equity and quality is disappointing (Barnett et al., 2010; Naslund-Hadley et al., 2013). From 1992 issues such as low enrolment, high dropout, high repetition and low completion have plagued secondary education ((Barnett et al., 2010). Only 45% of secondary school age children attend secondary schools which is below the region's rate of 80% (Naslund-Hadley et al., 2013). Even though strides have been made in the last five years, Naslund-Hadley et al. (2013) reported that only 27% of children between the age of 13-16 complete secondary school on time, many enrolled will dropout or repeat before graduating, and many are not receiving satisfactory levels of performance on examinations. According to the Ministry of Education's (2015) latest statistic, the secondary level has a net enrolment rate of 52.1%, a repetition rate of 7.2% and a dropout rate of 8.3%. The latest completion rate reported in 2014 was 61.1% which is a decrease from 2013. Furthermore, Naslund-Hadley et al.'s (2013) evaluation of Belize's education system stated that many of the teachers incorporate outdated pedagogical models and use materials that are predominantly teacher-centre. In addition, there is no accountability for how resources are used, no consequences for underperforming schools and low parental involvement in schools.

Having identified these issues, the MOE's educational plan for 2011 – 2016 set out to improve these ills. The sector plan proposes several strategies such as standardising secondary curriculum, developing series of desired learning outcome for all stages, establishing a National Qualifications Framework and developing and introducing a standardised National Secondary School Certificate (MOE, 2012). The plan also aims at increasing the quality of education by creating competent and committed teachers. One of the major strategy to fulfil this aim is to train teachers in learning assessments with emphasis on formative assessments to reduce failure and avoid repetition (MOE, 2012). The MOE also plans to establish minimum professional standards for teachers' education programs across the country.

2.5 SUMMARY

This section looked at the context of Belize including the geographic, political, economic and social features. It also provided an insight into Belize's education system, specifically as it relates to secondary education. The need for a sound and viable educational system is of tremendous importance in order for Belize to tackle its problems. A major way to improve the quality of education is through the proper usage of assessment. Assessment could be used as a major strategy to improve the dropout, repetition and completion rates at secondary level. However, the assessment of secondary school students needs to be nationalized and aligned with national educational standards.

CHAPTER THREE
3.0 LITERATURE REVIEW

3.1 INRODUCTION

This section will discuss the relevant literature pertaining to teachers' assessment practices and their conception of assessment. The literature will show that not much has changed with the usage and view of assessment. Teachers' conceptions and practices regarding classroom assessment are still quite similar to their predecessors thirty years ago. This is despite countless amounts of literature and research showing and advocating the benefits of alternative assessment methods. With this in mind, the review starts with an explanation of the theoretical framework encompassing this study. This is followed by sections that frame the issues of the study. These include an overview of the historical context and tension within the field of assessment, defining classroom assessment and defining assessment conceptions. The subsequent sections provide detailed analysis of the major themes and empirical evidences regarding assessment conceptions and practices. This is followed by an analysis and discussion of the relationship between conceptions and practices. The last section deals with teachers' assessment practices and conception in Belize. The review then concludes by summarizing the main points which illustrate the need for the study.

3.2 THEORETICAL FRAMEWORK

Classroom assessment sits at the intersection of theory and practice and it integrates instruction, management and assessment (Shepard, 2013; Brookhart, 2004). This makes it difficult to place classroom assessment within a specific framework since there are various theories that guide these areas. It is even more difficult to place this research within a specific framework since it looks at both practices and conceptions. Nonetheless, since classroom assessment is the core of the research, theories within this field will be used to create the theoretical framework.

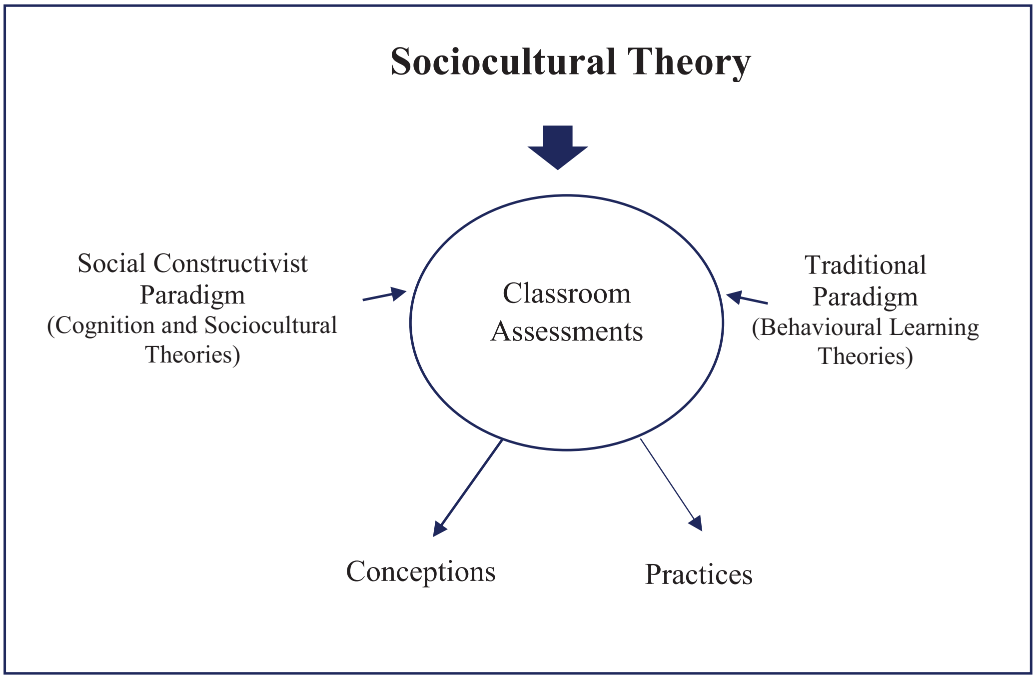

Figure 3.1. *Theoretical Framework of the Study*

Figure 3.1. illustrates the theories that the research is based upon. Sociocultural theory is the overarching theory of the research because it is able to account for mental processes, practices and contexts (Allahyar & Nazari, 2012; Cross, 2010). The units of analysis in this study are teachers who are considered social agents that act or react to stimuli in the educational context (Cross, 2010). Sociocultural framework acknowledges that teachers' cognition is dynamic and is greatly dependent on historical experiences and the social context of which they work (Cross, 2010; John-Steiner & Mahn, 1996). Thus, it provides a framework to understand teachers' assessment practices and conceptions within an educational context.

Sociocultural theory was first organized and used by Lev Semanovich Vygotsky from Russia in the nineteenth twenties and thirties (John-Steiner & Mahn, 1996). Vygotsky believed that "human activities take place in cultural contexts, are mediated by language and other symbol systems, and can be best understood when investigated in their historical development" (John-Steiner & Mahn, 1996, p.191). This theory has since greatly influenced the field of teaching and assessment.

However, only two principles of the theory that create the basis for the examination of teachers' thinking and practices will be discussed in this section. The first principle is that learning is dependent on both individual and social processes. A person acquires knowledge through social processes and then uses the acquired knowledge to form their own individual knowledge. This principle is in line with the main idea in this research because it is believed that societal processes such as the influence of former teachers, professional education and local educational context all impact teachers' conceptions of assessment which influence their assessment practices (Alkharusi, 2010; Shepard, 2000; Skott, 2015; Young & Kim, 2010)

The second principle is that the individual and social processes are mediated by tools and signs (Cross, 2010; John-Steiner, 1996). Tools such as language, mnemonic techniques and even the computer help to facilitate the co-construction of knowledge and help to regulate the world (Cross, 2010). These tools; however, are not created by an individual in isolation, but are the products of an individual's culture, history and social context (Allahyar & Nazari, 2012). Thus, Vgotsky believed that the best way to examine mental processes is through genetic or development analysis (Cross, 2010; John-Steiner, 1996). This entails studying the individual's mental processes as they develop within a particular culture, within the individual's life span and within the societal and culturally shaped context (Cross, 2010). This applies to this study in several ways. First, teachers' assessment conceptions and practices will be studied within the educational context of Belize. Using empirical evidence from other countries, it will be determined if Belizean teachers' practices and conceptions are different or similar to that of teachers in other countries. Second, teachers' demographic information such as teaching experiences and gender will be examined in relation to their conceptions and practices. Lastly, the study will examine the relationship between teachers' conceptions of assessment and their assessment practices to determine if what teachers believe about assessment is what they practice.

While sociocultural theory is the overarching theory of this research, it has also influenced the field of assessment along with other theories. Behaviourist learning theories formed the basis of the traditional paradigm in assessment (Dayal &

Lingam, 2015). These theories are mainly interested in the amount of knowledge students have acquired (Gipps, 1994; Serafini, 2000). Knowledge in this sense is seen as being separated from the learner and something that can be measured objectively (Serafini, 2000) and the main purpose of assessment is reporting. Social constructivist paradigm, on the other hand, is based on cognition and sociocultural theories which are interested in the learning process and context (Shepard, 2000). Here, knowledge is created by both the learner and teacher within a social and cultural context and the main purpose of assessment is not to report the amount of knowledge acquire, but to inquiry about the student's learning and learning progress (Serafini, 2000). Both paradigms and relating theories are discussed in more details in the subsequent section.

3.3 HISTORICAL OVERVIEW OF ASSESSMENT

The concept of assessing students' knowledge and achievement is not new (Harlen, 2005; O'Connor, 2007; Shepard, 2000; Stiggins, 2002; Ward and Murray-Ward, 1999). According to Ward and Murray-Ward (1999) assessment could be dated back to China nearly 4000 years ago where performance tests were used to choose and promote public officers. European universities also used assessment in the form of oral examinations to categorize students. It was not until the nineteenth century when psychologists started to study human intelligence that standardized testing and intelligence testing became dominant.

3.3.1. Behaviouristic and Psychometric Models of Teaching and Testing

Behaviouristic model of teaching and testing was dominant in the early twentieth century. It conceptualized teaching and testing under key assumptions such as learning occurs by collecting bits of knowledge, learning is sequential and hierarchical and testing is done frequently to ensure mastery (Shepard, 2000). Behaviouristic model reduces the curriculum to specific learning goals which are learned via structured, teacher directed methods. The frequent testing feature of behaviourist model gave rise to psychometric testing which was grounded in the belief that intelligence could be measured and, based on an individual's performance, that individual could be streamed and grouped within a school in order to enhance achievement (Berry and Adamson, 2011; Gipps, 1994; Shepard, 2000). Norm-reference was emphasized over criterion reference; thus, students' scores were mainly analysed in relation to their peers (Gipps, 1994). Behaviouristic and psychometric models of teaching and testing were common until researchers started to question their effectiveness in capturing students' achievement, much less improving students' learning (Gips, 1994; Shapard, 2000; Stiggins, 2002; Stiggins, 2014; Ward & Murray-Ward, 1999).

The main argument brought forth against behaviouristic and psychometric models was that it did not support the learning process (Gipps, 1994; Glasser, 1990; Gullickson, 1985; Resnick, 1989; Stiggins 2002; Stiggins, 2014). These models mainly assess isolated facts which were easily forgotten by students after they had taken a test; thus, true learning did not really occur (Gipps, 1994; Glasser, 1990). It has now been established that learning occurs best when students are provided with opportunities to interpret the content rather than just memorizing it (Resnick, 1989). Assessment, then, became a tool that helped with the learning process rather than merely assessing whether learning goals have been achieved. It started to stand in "dynamic interaction" with teaching and learning rather than standing outside of it (Gipps, 1994, p. 15).

3.3.2. Cognitive and Sociocultural Theories

This new paradigm of assessment was mainly based on the works of two influential theorists: Jean Piaget and Vygotsky. Piaget's development of cognitive learning theory and Vygotsky's development of sociocultural theory are credited for the shift in teaching and assessment (Berry and Adamson, 2011; Gipps, 1999; Shepard, 2000). Cognitive learning theory is concerned with the internal mental process of a learner (Krause et al., 2010). The learner mentally goes through the process of transferring, coding, storing and retrieving new information. Constructivist theory extended this concept by putting forth the idea that cognition does not occur on its own, but it occurs in collaboration with self-reflection, social processes and the environment

(Krause et al., 2010). Constructivism maintains that in order for learners to make sense of new learning, they must first construct their own learning by doing their own interpretation and organization of the new information (Shepard, 2000). In this sense, knowledge is neither passively received nor mechanically reinforced, it is an active process of making sense of new information (Gipps, 1999; Shepard, 2013).

While Piaget was concerned with the universal cognitive development of a child, Vygotsky was concerned with the differences in a child's cognitive development based on social, cultural and historical contexts (Krause et al., 2010). While it is accepted that learners construct their own knowledge by making meaning of the new information and then mapping it into their existing knowledge schema (Gipps, 1999), Vygotsky believed that these cognitive developments occur through social interactions (Krause et al., 2010). He termed the distance between the learner current competency level and the level they can achieve with social support and guidance as the zone of proximal development (Krause et al., 2010). A teacher then is responsible for guiding a learner through social activities that stretches their understanding of a skill which will eventually lead to the reduction of the gap between current learning level and desired learning level (Gipps, 1999; Shepard, 2013). Vygotsky's sociocultural theory contributed to the development of social constructivism which simply asserts that social processes are fundamental to learning (Krause et al., 2010). Combined together, cognition learning theory and sociocultural theory maintain that knowledge is not simply something that is stored within an individual, but knowledge is constructed in a learner's mind through social interactions.

3.3.3. Relation to Teaching and Learning

This new outlook of knowledge construction changed teaching and learning within the classroom. Teaching and learning became more interactive, dynamic and collaborative. Students became active learners who made sense of information for themselves with assistance and guidance from their teachers. Teaching became learner centred whereby opportunities were created for learners to work together and learners were empowered to become experts in their own learning creation (Krause et al., 2010). Assessment became vital in this new paradigm and sat at the intersection between teaching and learning. It became more than teachers giving tests and recording grades. Teachers started to assess students' learning by providing detailed information regarding students' understanding and providing purposeful feedback (Shepard, 2000). Alternative approaches to assessment such as performance assessment and authentic assessment also came to the forefront (Gipps, 1999).

3.4 TENSION AND CONFUSION BETWEEN PARADIGMS

The shift from a behaviouristic / psychometric model to a social constructivist model is filled with tension and confusion (Brookhart, 2004; Shepard, 2000). This is mainly due to the different and conflicting theories the two paradigms are based upon. Shepard (2000) illustrated the tension and confusion using the following figure.

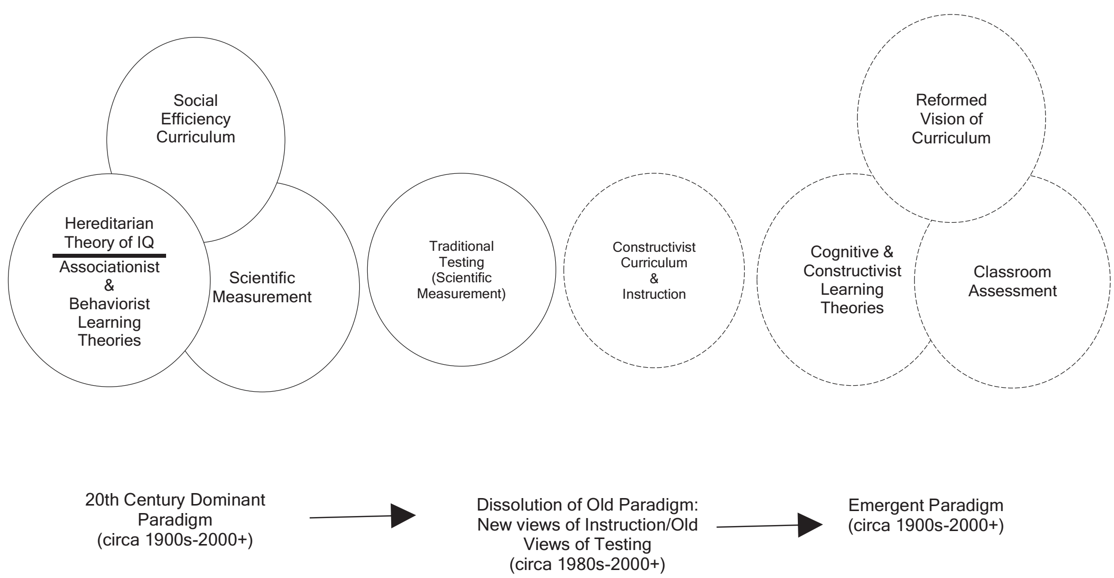

Figure 3.2. *An historical overview illustrating how changing conceptions of curriculum, learning theory, and measurement explain the current incompatibility between new views of instruction and traditional views of testing.* Adapted from "The Role of Classroom Assessment in Teaching and Learning," by L.A. Shepard, 2000, *Educational researcher, 29*(7), 4-14.

The left side of the figure shows the three main theories that have dominated teaching and assessment in the twentieth century. These theories gave rise to the psychometric model of testing and traditional testing. The right side shows the emergent theories that have given rise to alternative methods of assessment and formative assessments. The middle depicts a paradigm that has separated instruction from assessment. It was noted earlier that cognition and constructivist learning theories (representing the emergent paradigm) place assessment within, rather than outside, the teaching and learning dynamic. The middle of Shepard's illustration shows that this is not the case; assessment is in fact separated from instruction, and here lies one of the cause for tension and confusion. A shift from the dominant behaviouristic \ psychometric model to a cognition / social constructivist model means that assessment should be placed within teaching and learning. The fact that assessment still stands outside of instruction makes one questions if a shift in paradigms has even really occurred.

Beside the differences in theories, the dominance of traditional theories has also created tension and confusion within teaching and assessment (Berry & Adamson, 2011; Shepard, 2000; Shepard, 2013). One reason for the continuing dominance of traditional theories is that it has infiltrated into cultural beliefs (Shepard, 2000). Society, and teachers in particular, have come to regard the traditional form of teaching and testing as the norm. Thus, a teacher may see a feature of traditional testing such as norm-references or intelligence grouping as a normal part of teaching and testing and because of this belief, he or she may be resistant or hesitant in using emergent practices. Cultural acceptance is also made clear by the fact that researchers have discovered that many teachers utilize teaching and assessment practices used by their predecessors making the influence of traditional theories "potent but invisible long after they are abandoned by theorists" (Shepard, 2000, p.4). Knowing and understand teachers' assessment conceptions is more significant than knowing and understanding their practices. This is why educational assessment specialists have repeatedly advocated for studies on teachers' beliefs because any change in teachers' assessment practices will only come about when teachers' assessment beliefs are known and understood (Shepard, 2000; Stiggins 2002).

3.5 DEFINING CLASSROOM ASSESSMENT

Classroom assessment is an extensive and changing field involving both student and teacher (Martell, 2007; McMillan, 2013; Pittaway & Edwards, 2012). It has been defined in various ways by different educational specialists over the years. For instance, Harlen (1994) defined assessment as a process of gathering evidence and interpreting that evidence in light of some defined criterion to form a judgement. While Black and Wiliam (1998) defined it as all the activities undertaken by teachers and students. Gredler (1999) and Chase (1999) refer to classroom assessments as systematic procedures and methods that are used to gather information about students' learning. Madusa (1995) mentioned that whatever word is used; assessment, exhibition, examination, portfolio, test, it simply means that a sample of behaviours from a specific domain is used to make references about students' performance within that domain (as cited in Broadfoot, 2007). McMillan (2013) synthesized all these definitions by providing the following characteristics when defining classroom assessment. He stated that classroom assessment is

- used for collection and evaluation of students' learning
- done for varied reasons including diagnosis and monitoring
- a tool used by teachers to gather relevant data and information to make references about what students know, understands and can do
- vehicle through which student learning and motivation can be enhanced
- different from large-scale or standardization tests
- locally control
- structured techniques such as test, self-assessment and informal techniques such as observation and questioning
- more than mere measurement or quantification of students' performance (p. 4).

From the list of characteristics, we see that classroom assessment is the usage of different tools to collect information about students' learning that is then used to evaluate and monitor students' learning. Classroom assessment enhances students' learning and motivates students. It is not the same as large-scale or standardized test, it is not controlled by an external body and it is not solely a measurement or a quantification of students' learning. This research will specifically be looking at classroom assessment as defined by Zhang and Burry-Stock (2003) as a broad variety of activities ranging from constructing paper-pencil test and performance measures to grading, communicating test results, and using assessment results.

Formative and summative are the two main purposes of classroom assessment (Torrance & Pryor, 1998; Cizek, 2013). When used formatively, assessment guides and improves students' learning and teachers' instruction (Cizek, 2013). This usage is in line with the emergent paradigm explained earlier that is based in cognition and sociocultural theories. In contrast, when assessment is used in a summative way, it has a reporting function whereby inferences are drawn from evidence gathered from assessment practices to report students' learning to an audience for consequential purposes (Cizek, 2013). This type of usage is in line with the traditional paradigm based on behaviourist and measurement theories. Other purposes such as enhancing students' confidence and motivation and modelling achievement targets have also been identified (Cizek, 2013; Stiggins, 2014; Ohlsen, 2007). This research is concerned mostly with formative and summative usages which are discussed in detail later in the review.

3.6 DEFINING TEACHERS' CONCEPTIONS ABOUT ASSESSMENT

The difficulty in studying teachers' belief lies with the word choice, definition and eliciting those beliefs (Erkmen, 2012; Fives & Gill, 2015; Pajares, 1992; Skott, 2015). This difficulty mainly has to do with the understanding given to beliefs and knowledge. Pajares (1992) ascertained that "belief is based on evaluation and judgement; knowledge is based on objective facts" (p. 313). Kagan (1992) added by explaining that belief is better understood based upon teachers' professional knowledge. He stated that knowledge is belief that has been endorsed as truth based on the collection of opinions and objective proofs. Abelson (1979) presented seven features that distinguish belief from knowledge.

1. The elements (conceptions, positions, rules) of a belief system are not consensual.

2. Belief systems are in part concerned with the existence or nonexistence of certain conceptual entities.
3. Belief systems often include representations of “alternative worlds”.
4. Belief systems rely heavily on evaluative and affective components.
5. Beliefs system are likely to include a substantial amount of episodic material.
6. The content set to be included in a belief system is usually highly “open”.
7. Beliefs can be held with varying degrees of certitude.

(p. 356 – 360)

Nespor (1987), using Abelson’s (1979) framework, developed a conceptual framework for the analysis of belief. He explained that belief reflects the supposition about the existence or nonexistence of different types of entity, is an ideal situation which differs from the present, is organized around personal, cultural or political experiences, and is guided by feelings and moods. He stated that “to understand teaching from teachers’ perspectives we have to understand the beliefs with which they define their work” (p. 323).

As researchers started to study teachers’ overall perceptions and awareness of assessment, two terminologies arose; conception and value. Both terminologies submerged knowledge and belief into one construct and became the words that defined teachers’ beliefs regarding assessment (Barnes et al., 2015). Pratt (1992) likened conception to a lens that teachers use to filter their view, interpretation and understanding of the world. He defined conception as

> specific meaning attached to the phenomenon which then mediate our response to situation involving those phenomenon. We form conception of every aspect of our perceived world, and in so doing, use those abstract representations to delimit something from, and relate it to, other aspect of our world. In effect, we view the world through the lenses of our conceptions, interpreting and acting in according with our understanding of the world (p. 204).

Thompson (1992) saw conceptions “as a more general mental structure encompassing beliefs, meaning, concepts, propositions, rules, mental images, preferences and the like” (p. 130). Brown (2006) used the word conceptions to describe teachers’ perceptions toward assessment. He saw teachers’ conception as a reflection of teachers’ thinking about the purpose of student assessment. He believed that conception acts as a frame of reference to how teachers use and interpret assessment policies and regulations. Others such as James and Pedder (2006) used the word value to describe teachers' assessment related beliefs, while Ndalichako (2015) used the word perception. This research will utilize the word conception as used by Brown (2006) to refer to the beliefs or perceptions that teachers have towards assessment.

3.7 FOUR CONCEPTIONS OF THE PURPOSE OF ASSESSMENT

As mentioned earlier, three main purposes of assessment are improving teaching and learning, certifying students' learning and making schools and teachers accountable (Torrance & Pryor, 1998). Brown (2004) developed an inventory of teachers' conceptions of the purpose of assessment by conceptualizing four purposes of assessment (Brown, 2003; Brown, 2004; Brown, 2006; Brown et al., 2009; Brown et al., 2011b; Brown & Harris, 2009; Remesal, 2009; Segers and Tillema, 2011). Brown (2003) used the commonly accepted purposes and added one other purpose to create four conceptions of the purposes of assessment namely; assessment improves teaching and learning, assessment makes student accountable for learning, assessment makes school accountable for learning and assessment is irrelevant to the life and work of teachers and students. (Brown, 2003; Brown, 2004; Brown, 2006).

3.7.1. Improvement Conception

The main purpose of assessment is to improve teaching and learning (Black & Wiliam, 1998; Clark, 2012; Gipps, 1994; Martell, 2007; OECD, 2005; Stiggins, 2014). Teachers who perceive assessments as improving teaching and learning engage in practices that describe and diagnose students' performance. Information from assessment is used to plan teaching and direct students' learning. However, in order to be effective, information gathered must be valid, reliable and accurate descriptions of students' performance (Brown, 2004). If right and valid information is not gathered, bad decisions will be made which will worsen rather than improve teaching and learning (Stiggins, 2014). Since it serves to improve teachers' instruction and students' learning, this conception relates strongly with the principles of formative assessment or assessment for learning (Brown, 2003; Segers & Tilleman, 2011).

Teachers who believed in the improvement conception use a wide range of techniques including informal teacher based strategies and formal assessment tools. They focus on the process of students' learning rather than the product. They identify and fix impediments to students' learning and capitalize on students' unknown strengths (Brown, 2004). The involvement of students in these processes greatly enhance the improvement of teaching and learning (Black & Wiliam, 1998), thus teachers involve students in the decision making and improvement of their learning.

Popham (2005) asserted that it is pointless to do assessment if it is not used to improve students' learning. This view is shared by many teachers since studies on teachers' conception of assessment revealed teachers agreeing with the improvement view. Majority of the studies done by Brown and others using Brown's Teachers' Conception of Assessment (TCoA) inventory showed teachers mostly agreeing that assessment is done to improve teaching and learning (Azis, 2015; Brown, 2004; Brown et al., 2009; Brown et al., 2011a; Brown et al., 2011b; Brown & Michaelides, 2011; Segers & Tillema, 2011). Using TCoA-III, Brown (2004) did a study of primary school teachers in New Zealand to identify their conceptions of assessment. He discovered that majority of New Zealand teachers agreed with the improvement conceptions. Similar findings were found among teachers in studies done in Australia, Dutch, Greece, Hong Kong, China and Spain that also used Brown's TCoA (Azis, 2015; Brown et al., 2009; Brown et al., 2011a; Brown et al., 2011b; Brown & Michaelides, 2011; Remesal, 2009; Segers & Tillema, 2011).

Ndalichako (2015) developed her own questionnaire to explore secondary teachers' perceptions of assessment in Tanzanian. 4160 teachers who participated in the marking of the Certificate of Secondary Education Exam (CSEE) completed the questionnaire. Teachers indicated their level of agreement or disagreement using a five points Likert scale. Half of the teachers (50.9%) indicated that the purpose of assessment is to make decisions relating to teaching and learning.

However, other studies found varying findings. DeLuca and Lam (2014) used a qualitative method to explore beginning teachers' conceptions on student assessment in Florida. Fifty seven beginning teachers wrote an essay on their understanding and approach to using assessment within a diverse teaching context. The results revealed that beginning teachers were unable to see the value assessments had on informing teachers' practice; especially teachers' ability to individualized instruction for different learning needs. Teachers mostly saw assessment as serving a reporting function rather than an improvement function. Using a similar method, Dayal and Lingam (2015) explored in-service and pre-service teachers' conceptions of assessment in Fiji and found that even though majority of the teachers associated assessment to a formative function, 75% of them defined assessment in relation to a summative function. Majority of the teachers that stated that assessment should be used to improve teaching and learning were in-service teachers with only three pre-service teachers listing formative functions. These results indicated that differences with regards to improvement conceptions between in-service and pre-service teachers exist.

Remesal (2011) developed her own conceptualization of assessment purposes. She conceptualized a continuum of assessment practices that saw pedagogical conceptions that dealt with the monitoring of teaching and learning on one end, and societal / accreditation that dealt with teachers' accountability and achievement certification on the other end. Mixed conceptions that blended pedagogic with accountability came between the two. Using a qualitative methods, she explored fifty primary and secondary Mathematics pre-service teachers' conception about the functions of assessment in Barcelona, Spain. Using the continuum of assessment practices, Remesal (2011) created four dimensions to conceptualized teachers' belief about assessment; effects on teaching, on learning, on students' accreditation of learning and on teachers' accountability. Teachers went through two stages of interviews; the first stage was semi-structured, interpretive interviews with fifty student-teachers and the second stage used an event-recall technique type of interview with some of the fifty student-teachers. The results showed that while the teachers agreed that assessment should be used to improve teaching and learning, they separated beliefs of assessment that affect teaching from those that affect learning. Similar findings were found in Segers and Tillema (2011) who reported that Dutch teachers separate improvement that affect teaching from those that affect students' performance. Thus, teachers may see improvement purposes as either improving teaching or improving learning.

3.7.2. Accountability Conception

Beside improvement, teachers may conceive assessment as an accountability tool. According to Brown (2003), accountability could be positive and negative and carry high or low consequences. A student may repeat a class due to poor assessment results or a school may be publicly embarrassed due to poor exam results. Both of these are examples of negative accountability with high-stake consequences. On the other hand, scholarships may be provided to students based on their assessment results or teachers may get an increase pay due to students' assessment results which are examples of positive accountability.

Assessment, then, could be used to make both student and school accountable for students' learning. Students become individually accountable for learning through their performance on assessments (Segers & Tilleman, 2011). Based on their performance, students are placed through high stakes consequences such as graduation, promotion and selection (Azis, 2012). Grading, scoring, criterion reference tests and award certification are all used to make students accountable for their learning (Azis, 2012). Teachers and their schools may also be accountable for assessment results. Results from external examinations may be used to report on schools' performance. Even though a high stake consequence may not be attached to results reported, the public may use the reported results to deem schools as either effective or ineffective in providing quality education which may impact the enrolment rate in schools deemed ineffective since parents might choose not to send their children there. The fact that assessments carry such consequences put it in line with the traditional / summative perspective of assessment (Brown and Hattie, 2009).

As with improvement conception, teachers also seem to agree with accountability conceptions. For instance, in Ndalichako (2015), 20% of the teachers agreed that the purpose of assessment was to prepare students for final exams, while 6% agreed that the purpose was to assign marks or grades. Teachers in DeLuca and Lam (2014) stated that assessment is done to communicate to and satisfy external audience. They also stated that classroom assessment can be used by principal to keep teacher accountable. Dyal and Lingam (2015) asked teachers to choose from between two views the role that assessment plays in teaching and learning, 14% of the teachers chose the view that see assessment as the master of teaching and learning.

Results between the two accountability conceptions vary based on studies that used Brown's TCoA inventory. Majority of the teachers in Cyprus, Hong Kong, China and Indonesia agreed with the school accountability conception (Azis, 2015; Brown et al., 2009; Brown & Michaelides, 2011; Li & Hui, 2007), while teachers in New Zealand and Australia disagreed (Brown, 2011; Brown et al., 2011b). The mean score for teachers in Indonesia was (m= 4.66) which indicated a high agreement with school accountability conception (Azis, 2015). With regards to the student accountability conception, teachers in New Zealand also disagreed (Brown, 2011), but teachers in Australia and Hong Kong agreed with the conception (Brown et al., 2009; Brown et al., 2011b).

Correlation analysis showed teachers associating accountability to improvement. Accountability conceptions were positively correlated with improvement conception in all the studies. This implies that teachers may generally believe that in order for improvement to occur within teaching and learning, schools and students must be held accountable. This was clearly seen in the study with Dutch teachers who did not separate student accountability from improvement, but saw the purpose of assessment as related to both summative and formative purpose. Segers and Tillema (2011) noted that in the Dutch setting classroom assessments serve both formative and summative purposes; thus, Dutch teachers did not separate the two. This is in contrast to the results from Australia. Even though Australian teachers positively correlated school accountability to improvement, no correlation was found between student accountability and improvement. Remesal (2011) found similar findings in Spain among student-teachers who separated assessments that affect teaching from those that affect learning and differentiated beliefs about accountability from accreditation of achievement. These suggest that teachers may not associate the assessment of students' learning to improvement. Brown et al. (2011b) suggested that this is because the teachers might have had experiences where assessments were separated from learning. Nonetheless, majority of the studies reveal positive correlation between student and school accountability conceptions suggesting that results from students' assessment are used to hold school accountable (Brown, 2011; Brown et al., 2011b; Li & Hui, 2007).

These differences in findings regarding accountability conceptions are attributed to the context of the individual country (Barnes et al., 2015; Brown et al., 2009; Brown et al., 2011a; Brown & Hattie, 2009). New Zealand, Australia and Netherlands are considered low-stakes accountability contexts because they require few compulsory national assessments and assessments are mainly done at the local or school levels (Barnes et al., 2015; Brown & Hattie, 2009). Countries like China and Hong Kong are considered high-stakes contexts because public examinations carry consequences for teachers and students (Barnes et al., 2015; Brown et al., 2009; Brown et al., 2011a). In the context of low-stake testing, teachers may perceive assessment as improving learning and rely heavily on formative assessment practices, while those in high-stake context may perceive assessment as improving students' performance and tend to rely on summative assessment practices (Brown et al., 2011a)

3.7.3. Irrelevancy Conception

The fourth purpose of assessment as put forth by Brown (2003) is irrelevancy. This conception conceptualizes assessment as being irrelevant to the work of teachers and life of students. Brown argued that teachers who view teaching and learning as being student-centred will perceive assessment as being bad for students' learning, since such teachers are interested in the development of the whole child rather than just his or her intellectual growth. He built the case for irrelevancy conception by stating that assessment measures a small domain of knowledge, is subjective, is not made use of and is used to control and regulate students' behaviour. If teachers experienced assessment is such ways, they may develop the view that assessment is irrelevant. They may believe that tests and examinations serve no purpose and that they do not need those processes to judge students' achievement. Thus, they reject assessment and may see it as separate from teaching and learning practices (Azis, 2012). Furthermore, teachers may also reject assessment because of the effects it has on their autonomy and professionalism, or they might have witnessed the inaccuracy of assessment and thus reject it (Brown, 2004).

While majority of studies saw teachers agreeing that assessment is important, some showed teachers believing otherwise. Teachers in Indonesia, New Zealand, Netherlands, Hong Kong, Fiji and the USA disagreed with irrelevancy conception of assessment (Azis, 2015; Brown, 2011; Brown et al., 2009; Calveric, 2010; Dyal & Lingam, 2015; Segers & Tillema, 2011). Indonesia and Hong Kong teachers recorded the highest level of disagreement suggesting that these teachers strongly believed that assessment is important in teaching and learning (Azis, 2015; Brown et al., 2009). While New Zealand teachers disagreed with the statements that assessment is bad and assessment is ignored, they moderately agreed with the statement that assessment is inaccurate making Brown (2011) conclude that irrelevance "has more of an independent effect, rather than negative effect" in New Zealand teachers' thinking about assessment (p.61). On the contrary, teachers in Australia, China, Cyprus and Spain agreed with the irrelevancy conception of assessment (Brown et al., 2011a; Brown et al., 2011b; Brown & Michaelides, 2011; Remesal, 2009). Australian teachers' level of agreement was below 'slightly agree'. Their agreement with irrelevancy was credited to inaccuracy in assessment and assessment playing no legitimate role in education (Brown et al., 2011b). Chinese teachers' agreement with irrelevancy was credited to error in measurement (Brown et al., 2011a).

Brown and Hattie (2009) had reasoned that teachers who agreed that assessment improves teaching and learning were unlikely to simultaneously agree with the irrelevancy conception. While this was true for some of studies such as the one done in New Zealand where irrelevancy was positively correlated with improvement

(Brown, 2011), the Cyprus study proved differently. Cypriot teachers simultaneously agreed with both improvement and irrelevancy conception which are logically contradictory conceptions (Brown & Michaelides, 2011). Brown and Michaelides (2011) concluded that although assessment simultaneously elicited positive and negative attitudes, Cypriot teachers considered assessment to be essential and valuable in teaching and learning. The teachers may just be aware of the misuses of assessment policies and practices. This is justified from findings in New Zealand, Australia and China where irrelevancy was positively associated with student accountability suggesting that teachers are aware of measurement errors regarding students' assessment (Brown, 2011; Brown et al., 2011a; Brown et al., 2011b).

Interesting findings regarding teachers' view of assessment were discovered in Nadalichako (2015). It seems that teachers in Tanzania strongly believe that assessment is used to motivate and manage students' behaviour. This is so because 93% of secondary teachers agreed that assessments make students concentrate their effort on learning, 89% agreed that it makes students attentive in classroom and 89% agreed that it makes students confident in final exams. Furthermore, the fact that 46% of teachers agreed that assessments increase workload of teachers, 37% agreed that it should be conducted based on teachers' convenience and 19% agreed that it interferes with teaching and learning seem to be indicating that Tanzanian teachers believe that assessment is a burden to teachers rather than an important tool.

3.7.4. Differences among Studies on Teachers' Assessment Conceptions

It is obvious that teachers' conception are multifaceted and interconnected (Barnes et al., 2015; Brown & Hattie, 2009). While majority of the teachers in the studies agreed that assessment is done to improve teaching and learning, they also believed that assessment holds students and schools accountable (Brown, 2004; Brown et al., 2009; Brown et al., 2011a; Brown et al., 2011b; Brown & Michalis, 2011; Remesal, 2009; Remesal, 2011; Segers & Tillema, 2011). Some teachers do believe that assessment should be ignored because it is irrelevant to teaching and learning or because it is inaccurate (Brown et al., 2011a; Brown & Michalis 2011; Remesal, 2009). Tensions between improvement for students' performance and improvement for teaching were also revealed. While Brown (2004) and Brown et al. (2011b) combined the two under one conception, other studies found that teachers necessarily do not view it that way (Remesal, 2011; Segers & Tillema; 2011). Segers and Tillema (2011) discovered that Dutch teachers did not see the purpose of assessment related to summative and formative purposes separately. In the same way, Remesal (2011) found that Spanish teachers generally separate beliefs of assessment that affect students' performance and those that affect teaching.

These differences along with differences revealed between low-stake and high-stake contexts (Brown et al., 2009; Brown et al., 2011a), signify the need for individual country's examinations of teachers' conception of assessment. Azis (2012) reviewed several studies on teachers' conception of assessment conducted in different countries and concluded that "The choice and deployment of [assessment] tools and the interpretation of the results depend on educational system, public and political influence" (p.46). Similar conclusions were made by Brown and Hattie (2009), Brown and Michaelides (2011) and Brown et al. (2011a). Brown and Michaelides (2011) pointed out that teachers' conceptions of assessment tend to be in line with the features of policies and practices within a particular country. For example, in Cyprus filing and using assessment results is not mandatory and schools are rarely monitored by external bodies. Thus, policies and monitoring of assessment rest with individual schools and principal's guidelines. This may lead to varied conceptions about assessment among teachers in Cyprus and may even lead to teachers not trusting assessment results as was reveal in Brown and Michaelides' (2011) study. Brown et al. (2011a) noted that the differences lie not only with culture and society, put also with adaption and implementation of policies. Countries, and even schools, may adapt and implement policies concerning assessments differently. For instance, even though Hong Kong and China share similar educational and cultural contexts, Chinese teachers agreed with irrelevancy conception while Hong Kong teachers disagreed (Brown et al., 2011a). This difference was believed to be due to an improvement oriented assessment policy in Hong Kong in contrast to a "high-pressure selection-orientation of examinations" in China (p. 315).

Variations in findings regarding the relationship between teachers' characteristics and their conceptions of assessment were also discovered. Brown (2004) reported no significant differences in mean scores between teachers' years of experience and measurement training and assessment conceptions. He concluded that "differences in role are not powerful in shaping teachers' conception of assessment" (p. 311). However, Remesal (2009) reported differences among beginning teachers based on subject area. Brown and Gao (2015) also found differences with teachers' gender and teaching experience. They discovered that male teachers and teachers with twenty or more years of experience agreed more with the belief that assessment should be used to inspect and control school, teachers and students in order to promote better teaching and learning. Statistical differences were also found between teachers with more years of experience and the lack of agreement with belief that assessment should be done to diagnosis the effectiveness of teaching and direct teaching. Ndalichako (2015) also found that more female than male teachers have a favourable perception of assessment. She conducted a *t*-test and discovered a significant statistical difference between female and male teachers regarding the use of assessment to facilitate and support teaching.

Even though the studies showed that different methodologies could be utilized to study teachers' conceptions, the commonly used methodology is quantitative. Of the eleven studies presented on teachers' conception of assessment, nine used a quantitative method (Brown, 2004; Brown et al., 2009; Brown et al., 2011a; Brown et al., 2011b; Brown & Michalis, 2011; Ndalichako, 2015; Remesal, 2009; Remesal, 2011; Segers & Tillema, 2011) while two used a qualitative method (Dayal & Lingam, 2015; DeLuca & Lam, 2014). Even though many of studies used a version of Brown's TCoA inventory, caution should be taken when using it because some of the studies revealed a misfit with their context of teachers (Brown et al., 2009; Brown et al., 2011a; Brown & Michalis, 2011; Remesal, 2009; Segers & Tillema, 2011). Teachers who participated in the studies came mainly from the secondary level and were practicing teachers with the exceptions of Remesal (2009) who used freshman student-teachers and DeLuca and Lam (2014) who used beginning teachers.

3.8 FORMATIVE ASSESSMENT / ASSESSMENT FOR LEARNING

For a long time attention was placed on assessing students' learning at the end of a unit of activities, with little attention given to activities leading up to that end (Crossouard & Pryor, 2012; Wiliam, 2013). Over the past twenty years researchers have started to investigate and examine activities done in the classroom and have ascertained that specific classroom activities have the power to improve students' learning and achievement. This attention placed formative assessment at the centre of teaching and learning improvement. Formative assessment has attracted a great deal of attention because it acts as an intersection between curriculum and assessment, it supports learning rather than accredits learning, it is encoded into teaching and it allows teachers to direct students' learning (Crossouard & Pryor, 2012).

There is no clear consensus on the meaning of formative assessment (FA) and assessment for learning (AfL) (Clark, 2010; Wiliam et al., 2011). Black and Wiliam (1998) use the general term assessment to define formative assessment as:

> …all those activities undertaken by teachers – and by their students in assessing themselves – that provide information to be used as feedback to modify teaching and learning activities. Such assessment becomes formative assessment when the evidence is actually used to adapt the teaching to meet student needs (p. 140).

Kahl (2005) defined FA as "a tool that teachers use to measure student grasp of specific topics and skills they are teaching. It's a 'midstream' tool to identify specific student misconceptions and mistakes while the materials is being taught" (p.11, as cited in Wiliam, 2011, p. 10). OECD (2005) stated that FA refers "to frequent, interactive assessment of student progress and understanding to identify learning needs and adjust teaching appropriately" (p. 21).

Others such as The Assessment Reform Group (a group that ensure assessment policy and practices are informed by research evidence) favours the term assessment for learning instead of formative assessment (Wiliam, 2011). They argued that formative assessment only constitutes the frequent usage of assessment as a part of teaching, which does not necessarily mean that such assessment is helping to advance students' learning. Believing that the term "formative" is open to various interpretations, they prefer the terminology assessment for learning and defined it as "the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there" (Broadfoot et al., 2002, p. 2-3). The distinction between FA and AfL is further shown by Black, Harrison and Lee, (2004) who distinguished the two by stating that AfL prioritizes on promoting student learning through the use of assessment information, but such assessment only becomes "formative" if the information is indeed used "to adapt the teaching work to meet learning needs" (p.10).

Many authors use formative assessment and assessment for learning interchangeably. For instance, James and Pedder (2006) used FA and AfL as equivalent and interchangeable in their research on teachers' perception of classroom assessment practices. Wiliam (2013) asserted that the importance was not the distinction between FA and AfL, but the distinction between formative and summative assessment. Thus, both will be used interchangeable within this research since both refer to assessments that are conducted with the intention to improve learning and evidence from the assessments are used to improve instruction (Wiliam, 2013).

Formative assessment is often thought of as gaining evidence about the gap between students' actual learning level and the desired learning level, but it is far more than gathering evidence about a gap (Black & Wiliam, 2009; Wiliam, 2013). Before the evidence gathered is used, it must be verified that such information is valid and credible because inaccurate information will just lead to bad decision making regarding teaching and learning (Stiggins, 2014).

The number one way to use evidence gathered to improve students' learning is to provide students with feedback. Feedback is the information about the gap in students' learning (Hattie & Timperley, 2007; Hughes, 2010; Sadler, 1989). Feedback is only effective if it is of good quality and if it is used to make changes in teaching and learning. Sadler (1989) outlined three conditions that should be utilized to make feedback effective. The student must know the standard being aimed for, know the difference between his or her current performance and the desired performance level and actions should be taken to reduce or eliminate the gap. Such strategies have been claimed to greatly maximize the learning of students especially those that are considered at-risk (Black & Wiliam, 1998; Cizek, 2013; Hughes, 2010). At- risk students lack the environment to support their needs (Huges, 2010) and high quality feedback helps them to develop skills such as self-monitoring, metacognition and self-evaluation which are then used to create the proper environment (Huges, 2010). Hattie and Timperley (2007) extended Sadler's (1989) conditions by providing three questions that feedback should answer: where I am going, how I am going and where next. Answers to these questions work on four levels:

1. Feedback about task
2. Feedback about the processing of the task
3. Feedback about self-regulation and
4. Feedback about self as a person (as cited in Wiliam, 2013, p. 204).

Going through these processes creates effective feedback that could be used to improve students' learning.

Formative assessment also requires learners to be engaged in actions to improve learning such as seeking assistant from peer and consistently reflecting on ways to enhance their learning (Black & Wiliam, 2009). According to Topping (2013), "Peer assessment is generally an arrangement for classmates to consider the level, value, or worth of the products or outcomes of learning of their equal-status peer" (p. 395). Peer assessment has both long and short term benefits and it assists both the person doing the assessment and the person being assessed (Black & Wiliam, 2011; Topping, 2013). It acts formatively by informing and redirecting students' effort and connecting outputs with objectives, while it acts as a summative by providing only numerical or categorical data (Topping, 2013). Self-assessment; on the other hand, is students reflecting on their own learning. Brown and Harris (2013) defined it as a "descriptive and evaluative act carried out by the student concerning his or her own work and academic abilities" (p. 368). Both self-assessment and peer assessment have been shown to be vital in enhancing students' learning.

Noonan and Duncan (2005) discovered that a large portion (76%) of high school teachers in Western Canada used some type of peer and self - assessment when assessing students' learning. When asked why they used such practices, teachers stated that the practices encourage and facilitate students' reflection on their achievement and it is useful in assessing group works, projects and presentations. James and Pedder (2006) also found that majority of the teachers in their study stated that they highly valued such assessment practices because it provide opportunities for students to assess their work and learning. Brown and Harris (2013) did a study that examined studies done on students' self-assessment and concluded that there is empirical evidence that self-assessment of a task equals to improve academic performance.

A comprehensive list of ten characteristics that describes formative assessment was compiled by Shepard (2006). According to Cizek (2010), this list (Table 1) was compiled based on the work of Black and Wiliam (1998), Sadler (1989) and Atkins, Black and Coffery (2001). This list places formative assessment as a tool that could be used to guide students' learning and to provide teachers with information that can be to plan instruction (Cizek, 2010).

1. Requires students to take responsibility for their own learning.
2. Communicates clear, specific goals.
3. Focuses on goals that represent valuable educational outcomes with applicability beyond the learning context.
4. Identifies the student's current knowledge / skills and the necessary steps for reaching the desired goals.
5. Requires development of plan for attaining the desired goals.
6. Encourage students to self-monitor progress towards the learning goals.
7. Provide examples of learning goals including, when relevant, the specific grading criteria or rubric that will be used to evaluate the student's work.
8. Provides frequent assessment, including peer and student self-assessment and assessment embedded within learning activities.
9. Includes feedback that is non-evaluative, specific, timely, related to the learning goals, and provides opportunities for the students to revise and improve work products and deepen understandings.
10. Promotes metacognition and reflection by students on their work.

Table 3.1: *Ten Characteristics of Formative Assessment* Adapted from "An Introduction to Formative Assessment," Cizek, G. (2010). In H. Andrade & G. Cizek, *Handbook of formative assessment* (1st ed., pp. 3-17). New York: Routledge.

Growing evidence has proven that formative assessment practices have the potential to improve students' learning and achievement (Black & Wiliam, 1998; Crossouard & Pryor, 2012; James & Pedder, 2006; Wiliam, 2013; William, Lee, Harrison & Black, 2010). Wiliam et al. (2010) conducted a study with twenty four teachers within six secondary schools to investigate the effects that increase use of formative assessment will have on students' achievement on standardized test. The range of formative assessment strategies introduced by the teachers were calculated to have an effect size of 0.32. Wiliam et al. (2010) estimated there to be a 70% increase in students' learning due to the formative assessment practices incorporated by the teachers. This collaborates earlier findings by Black and Wiliam (1998) who did a meta-analysis of 250 studies addressing various aspects of FA and discovered effect size between 0.4 – 0.7. On the contrary, Kingston and Nash (2011) conducted a reviewed of 300 studies that address the efficacy of FA and concluded that the effect sizes being claimed are not supported by existing research base. McMillan, Venable and Varier (2013) counteracted their findings by doing an analysis of Kingston and Nash's (2011) study. They found that there were several weaknesses in the methodology used and several limitations in the 300 studies analysed. McMillan et al. (2013) concluded by stating that the conclusions came to by Kingston and Nash

(2011) regarding formative assessment and achievement effect are tentative and implausible. While there is a need for more sound research on the efficacy of FA, it is still widely believed that formative assessment is able to improve students' learning and achievement.

Several studies have examined the usage of formative assessment practices among teachers (James and Pedder, 2006; McMillan et al., 2010; Noonan and Duncan, 2005; Volante and Beckett, 2011). Through the use of individual interviews, Volante and Beckett (2011) examined twenty elementary and high school teachers' formative assessment practices in Ontario, Canada. They reported that teachers are becoming familiar with formative assessment strategies such as questioning, feedback without grades, self-assessment, peer assessment and formative use of summative assessment. The study also revealed that teachers use such practices on a consistent basis. For instance, when asked about the use of feedback without grade, a high school teach replied that she allowed students to resubmit their work several times before the due date while providing seven different feedbacks to them. On the same note, to investigate teacher's formative assessment practices, McMillan et al., (2010) used teachers' reported formative assessment practices and related them to their students' self-reported motivation. High school teachers from Richmond Virginia and their students completed two self-reported questionnaire. Majority of the teachers (70%) reported using some form of formative strategies in monitoring students' progress; yet, only 23 % reported using assessment to diagnosis student weakness and 32% reported using it to guide future instructions. When teachers' practices were correlated with students' level of motivation, there was a slight statistical significance which indicated that some of the teachers' practices are important. Moreover, Song and Koh (2010) examined teachers' beliefs about student learning and formative assessment practices in Singapore using a mixed method approach. Ninety eight teachers across three high schools completed two self-reported questionnaires which was followed by a semi-structured interview with one of the teacher. From the questionnaires, Song and Koh (2010) reported teachers who believed that students are active participants in their learning are more likely to believe that students need to monitor their own learning. These teachers also held the belief that students learn by constructing their own knowledge. The most commonly used formative assessment practice was questioning. These results show the popularity in the usage of formative practices.

Despite that, it has been claimed that many teachers do not necessarily incorporate formative assessment practices in their classes (Black and Wiliam, 1998; Clark, 2012). Heritage (2007) stated that

> after more than hundred years of exhortations and a significant body of research on the topic, the idea that assessment and teaching are reciprocal activities is still not firmly situated in the practice of educators. Instead, assessment is often viewed in competition with teaching, rather than as an integral part of teaching and learning (p. 11).

Even though teachers have recorded the usage of formative strategies, in practice, they continue to see assessment as being outside of teaching and thus continue to overemphasize summative assessment practices such as tests and quizzes (Clark, 2012; Volante & Beckett, 2011). For instance, even though the teachers in Song and Koh's (2010) study reported that students learn by knowledge construction, the use of formal and informal summative assessment practices overshadowed teachers' formative assessment practices. Further, the study reported a lack of correlation between teachers' belief that students are active participants in learning and their usage of peer and self-assessments. This indicates that even though teachers believe in the principles of formative assessment, they are not necessarily using formative practices in their classroom.

3.9 SUMMATIVE ASSESSMENT / ASSESSMENT OF LEARNING

Summative assessments remain the common form of assessment used in the classroom. It is used for several purposes such as to measure students' learning at the end of a unit, summarize students' achievement at the end of a course of study, promote students from one level to another, measure standardization for certification, report to parent and students and entry into certain occupation and further education (Broadfoot & Black, 2004; Harlen, 2005; Looney, 2011; OECD, 2005; Sadler, 1989). It is not seen as a contributor to learning since it occurs at the end of a unit of study and does not give students the opportunity to improve (Looney, 2011; OECD, 2005).

Summative assessment practices are usually associated with traditional models of assessing students' learning (Gipps, 1999; Gredler, 1999; Harlen, 2005). In this sense, assessment is done outside of the regular instruction or at the end of instruction (Gipps, 1999; Gredler, 1999; Volante & Beckett, 2011). Teachers are the transmitters of information and assessment is done to check whether information has been received (Gipps, 1999; Gredler, 1999). Gredler (1999) demonstrated the typical instruction cycle with a maths class that utilizes a summative assessment model. The teacher firstly presents the concepts, then provides the procedures to understand the concept and lastly distributes worksheets for students to practice the concept. Volante and Beckett (2011) described such a model as "teach, test and hope for the best model" (p. 240). The purpose of assessment in such case is to merely categorize and select students (Gredler, 1999). Such purpose is also related to Brown's (2004) student accountability conception whereby students are held responsible for their learning through the use of grades.

However, researchers have argued that the number and nature of summative tasks could be used to improve students' learning (Black & Wiliam, 1998; Black et al., 2004; Harlen, 2005; OECD, 2005). Harlen (2005) asserted that summative assessment could be used to help students' learning by using the information from summative assessments to show the steps they need to take and ways they can take them. Black et al. (2004) reported on ways teachers in their study used summative results to help in the learning process. Teachers did so by helping students to prepare for test by reviewing past test papers to identify areas of weakness, asking student to set questions and marking schemes and using the results of test diagnostically.

3.10 TEACHERS' CLASSROOM ASSESSMENT PRACTICES

It has well been documented that teachers use a number and variety of assessment practices (Cizek et al., 1995; Popham, 2005). Even though alternative methods such as performance assessments, portfolio and authentic assessment have been considered as the ideal methods to assess students' learning over traditional methods of constructed response and selected response, those methods are rarely used in the classroom. An analysis of studies on teachers' assessment practice over the past thirty years will show that teachers' practices have rarely changed and have remained dominantly traditional.

3.10.1. Teachers' Assessment Practices in 1980s

Three studies done during the eighties showed similar results regarding teachers assessment practices. Through the use of a questionnaire, Stiggins and Bridgeford (1985) conducted a study with 228 elementary and secondary teachers across different subject areas in the U.S. to explore the nature and quality of teacher-developed assessments. The study revealed that majority of the teachers (85%) used some form of paper-pencil test with half reported that they relied mostly on published test. They concluded that teachers used a consistent amount of assessment practices across subjects and levels and this fact indicated that teachers did not change their practices to suit different purposes of assessment. They claimed that teachers might not have explored new assessment approaches due to the lack of opportunities, time, means or motivation.

A Similar study was conducted by Gullickson (1985) who surveyed 295 teachers across subject and grade levels to investigate their evaluation practices and to determine which were mostly used by the teachers. He also discovered that teachers used a variety of assessment practices, but highly rated practices such as teacher made test, discussion and paper / notebook. Like Stiggins and Bridgeford (1985), objective test was mostly used by the teachers, but classroom discussions and paper / notebook were also favoured. Gullickson (1985) concluded that secondary teachers relied on fewer techniques and used published test less often than primary teachers.

Marso and Pigge (1988) also looked at teachers' assessment practices but went deeper by analysing the quality and nature of teacher made test. One hundred and seventy five teachers completed a questionnaire on their assessment practices and submitted self-created tests. The results revealed that the most frequently used practices were multiple choices, matching and short response. Further, it was revealed that even though half of the teachers wrote their own test items, 80% never calculated test means and standard deviation, 6% never or rarely estimated reliability of test and 54% never or rarely completed item analysis. From the analysis of teachers' self-made test, it was revealed that majority of the items (72%) functioned at the knowledge level with only science and math teachers going beyond that level. Marso and Pigge (1988) concluded that the teachers seemed to value testing but were not convinced about the value of statistical procedures.

3.10.2. Teachers' Assessment Practices in the 1990s

Going into the 1990s, teachers' assessment practices continued to be dominated by summative traditional methods, but alternative methods based on formative purposes started to be revealed. Cizek et al., (1995) collected demographical information and assessment practices of 143 primary and secondary teachers from an urban Midwestern university. Similar results as Stiggin and Bridgeford (1985), Gullickson (1985) and Marso and Pigge (1998) were found; namely, that assessment practices varied and that majority of the teachers developed their own minor and major tests. Like Marso and Pigge (1988), Cizek et al. (1995) found that teachers developed their own test despite lacking the proper training. They concluded that "grades appear to consist of a potpourri of elements that vary across district to district, from teacher to teacher within a district" (p. 174).

Two similar studies done in 1998 on teachers' assessment practices were Bol et al. (1998) and Mertler (1998). Both used a survey to explore teachers' assessment practices. Bol et al. (1998) surveyed 893 primary and secondary teachers, while Mertler (1998) surveyed 625 kindergarten through to secondary teachers. Bol et al. (1998) discovered that teachers used practices such as observation, closed ended examinations, quizzes and assignments more frequently than other practices. This is similar to Mertler (1998) who found that teachers used summative practices slightly more than formative practices. In fact, the least used method reported in Bol et al. (1998) was self-assessment. Furthermore, both reported than teachers felt a little prepared in using the different assessment methods to assess students' learning. This low level of preparation could be seen in the fact that Metler (1998), like Marso and

Pigge (1988), found that a small amount of teachers (25%) actually followed steps to ensure validity and reliability.   From these two studies, it could be seen that teachers started to incorporate alternative methods of assessing students' learning even though traditional methods remained dominant.

3.10.3. Teachers' Assessment Practices between 2000 - 2010

The new millennium saw a surge in studies on assessment practices, but unfortunately, studies continued to reveal the dominant summative traditional assessment practices as prior studies. For instance, Zhang and Burry-Stock (2003) surveyed 845 teachers across school levels and subject areas in south eastern state in the U.S. to investigate teachers' assessment practices and their self-perceived assessment skills. The results showed similar findings as past studies such as the frequent usage of paper and pencil test and the heavy reliance on teacher made objective test.   Zhang and Burry-Stock (2003) concluded that assessment decisions based on grades took precedent over students' progress in the school system.

Three additional studies conducted in different time periods and contexts found similar practices among teachers. McMillan and Lawson (2001) examined the classroom assessment and grading practices of 261 secondary science teachers from sixty nine schools within the Virginia school district. Using a questionnaire, they found that most science teachers used a variety of factors in grading students such as effort ability, external benchmark, extra credit, academic achievement and participation. They also found that the four major types of assessment practices used were construct response, teacher-created assessment, objective assessment and major examination with objective examination being used more frequently than the others. On the same hand, Duncan and Noonan (2007), using an adapted version of McMillan and Lawson's (2001) questionnaire, examined teachers' assessment and grading practices in Western Canadian Province. Five hundred and thirteen secondary teachers completed the questionnaire and similar findings as McMillan and Lawson (2001) were discovered such as teachers frequently using objective assessments and frequently relying on students' efforts when grading. These findings were also echoed in Frey and Schmitt (2010) who surveyed 140 teaches across twenty two school districts in a mid-western state in the U.S. and found that the most commonly used practices were short answer, fill in the black and multiple choice. They also discovered that nearly half of the teachers used assessment created by themselves. Further, 87% reported that assessment used affected students' grades and 75% indicated that assessments were given after instructions.

3.10.4. Teachers' Assessment Practices in the Last Five Years

Similar findings among teachers regarding their assessment practices continue to be seen in studies conducted in the last five years. The one difference observed is that more studies are now available from countries other than the U.S. and Canada. For instance, Suah and Ong (2012) conducted a study with 406 teachers in the northern states of Peninsular Malaysia to identify their current assessment practices. Teachers were from different subject areas and school levels. Using a Teacher Assessment Practice Inventory created by the researcher, they found similar results as reported by other studies; namely that traditional type of assessment such as multiple choices, short answers and essay questions were highly used, while assessments such as project work and feedbacks regarding students' strengths and weaknesses were rarely used. They also found that content validity were not often done by the teachers a finding that has been constant among past studies.

Ogunkola and Clifford (2013) investigated the instructional assessment practices of 55 secondary science teachers in Barbados. They also found that teachers mainly used objective assessments such as tests and home assignments, but rarely used detailed feedbacks and portfolios. The use of objective tests and home assignment is validated with their main reasons for using assessments which were to determine understanding, assess cognitive skills and reinforce concepts taught.

Alkharusi conducted a series of studies with teachers in Oman. His studies did not only examine assessment practices of teachers, but examined how those practices affected students' learning. Alkarusi (2010) surveyed 83 science teachers and their students to examine how teachers' assessment practices impact students' perception. He found that classes have an assessment character that is influenced by the assessment practices used by the teachers. Teachers that used a low amount of recommended assessment practices produced low self-efficacy among their students. Such findings were corroborated in Alkarusi, Aldhafri, Alnabhani and Alkalbani, (2014) who found that teachers' high use of traditional methods had negative impact on students' self-efficacy beliefs. This study also found that teachers highly used non-achievement factors while grading students such as students' effort and ability. The interesting thing is that such factors correlated positively with students' self-efficacy beliefs. These findings show that assessment practices have direct and indirect effects on students' motivation making it very important to study.

3.10.5. Putting It All Together

Low usage of alternative practices based within formative purposes is revealed through the studies. This is despite the fact that such practices have been widely acclaimed as the best practices for improving student's learning. Alternative assessment practices require teachers' observation and professional judgement to make decision regarding student's achievement (Popham, 2005). It emphasizes knowledge construction and problem solving and encompasses several different assessment strategies including performance based, observation techniques, student self-assessments and portfolios (Bol et al., 1998). It is seen as the best approach because it promotes higher standard of teaching and more powerful learning. It goes beyond traditional assessment by motivating students to take more responsibilities for their learning, making assessment an integral part of learning and creating active learners rather than passive learners (Hargreaves, Earl & Schmidt, 2002).

Unfortunately, majority of the studies examined reveal that teachers rarely used such assessment practices, but seem to favour summative traditional practices (Gullickson, 1985; Frey & Schmitt, 2010; Marso & Pigge, 1988; McMillan & Lawson, 2001; Suah & Ong, 2012; Stiggin & Bridgeford, 1985). These practices include close-ended items, multiple choices, true and false, short-answer and fill in the blanks

(Gullickson 1985; Marso &Pigge, 1988; Stiggin & Bridgeford, 1985). It also includes using assessment only to determine students' overall achievement in order to promote, select or certify students. Such practices have been shown to have a negative impact on not only students, but on teacher and the curriculum (Moss, 2013). For instance, Halen and Crick (2002) discovered that after national curriculum tests were introduced in England, low-achieving students started to display lower self-esteem than high-achieving students. They claimed that before the tests were introduced, there existed no correlation between students' achievement and their self-esteem (as cited in Moss, 2013, p. 237). Students with low self-esteem tend to place less effort into future learning activities (Moss, 2013, Tierney, 2006). Summative traditional practices are usually adapted by teachers to suit testing. Such practices create a "performance ethos" (Moss, 2013, p. 237) in a classroom that portray learning as doing well on tests (Shepard, 2013). It is therefore very puzzling that after thirty years of studies on the benefits of using alternative methods based on formative purposes, summative traditional practices are still dominantly being used by teachers (Frey & Schmitt, 2010; McMillan & Lawson, 2001; Ogunkola & Clifford, 2013; Suah & Ong, 2012; Zhang and Burry-Stock, 2003).

It is also interesting that majority of the studies reveal inconsistent findings between teachers' assessment practices and subject area (Bol et al., 1998; Cizek et al., 1995; Duncan and Noonan, 2007; Frey & Schmitt, 2010; Mertler, 1998; Ogunkola & Clifford, 2013; Stiggins & Bridgeford, 1985; Zhang & Burry-Stock, 2003). Findings revealed that mathematics and science teachers' assessment practices differ from other teachers since they rely heavily on objective assessments (Duncan and Noonan, 2007; Frey & Schmitt, 2010; Ogunkola & Clifford, 2013; Ohlsen, 2007; Stiggins & Bridgeford, 1985). Mathematics and science teachers are thought to use paper and pencil tests, published tests, major exams and quizzes frequently than English Language and social studies teachers (Ogunkola & Clifford, 2013; Ohlsen, 2007; Stiggins & Bridgeford, 1985). Zhang and Burry-Stock (2003); however, revealed that such is not necessarily the case since they discovered that science, social studies and language arts teachers used paper and pencil test more often than mathematics teachers. Nevertheless, social studies and English language teachers were found to use performance assessments, essays, individual and team projects, oral assessments more often than science and English language teachers (Duncan and Noonan, 2007; Frey & Schmitt, 2010; Stiggins & Bridgeford, 1985). Mathematics and science teachers were also found to frequently use non-cognitive abilities as a contributor for grades and focused more on cognitive knowledge beyond knowledge level (Stiggins & Bridgeford, 1985; Duncan & Noonan, 2007).

McMillan (2013) explained such differences by stating that mathematics and science are usually taught and learnt from a predetermined sequence. This makes it easy for teachers to offer feedback and instructional correctives, but such practices are usually planned and structured. On the other hand, social studies and language teachers utilize student contributions and questioning more often than science and mathematics teachers in order to develop students' thinking. The predetermined and structured sequence allows mathematics and science teachers to gear more towards summative traditional practices. However, McMillan (2013) maintains that all subjects could be guided by constructivist learning theory whereby students are active participants in the teaching and learning process.

Studies also vary regarding differences between assessment practices and teaching experiences and professional qualification. Zhang and Burry-stock (2003) found significant differences between teachers' assessment practices and measurement training suggesting that training do have an impact on the practices that teachers utilize. However, Ogunkola and Clifford (2013) found no such differences between science teachers' assessment practices and their professional training. Similar inconsistent findings exist between teaching experiences and practices. For instance, Mertler (1998) found that novice teachers used alternative methods more frequently than experienced teachers, but Bol et al. (1998) found the opposite. They found that more experienced teachers used alternative methods more often than less experienced teachers. This was corroborated by Frey and Schmitt (2010) who also found that more experienced teachers used performance assessment more often than traditional assessment. Other studies found no significant differences between teaching experiences and assessment practices (Zhang & Burry-Stock, 2003; Ogunkola & Clifford, 2013).

3.11 RELATIONSHIP BETWEEN ASSESSMENT CONCEPTIONS AND PRACTICES

All the studies in this review emphasize the importance of examining teachers' conceptions toward assessment in order to better understand teachers' assessment practices (Barnes et al., 2015; Brown 2004; Ndalichako, 2015; Remesal, 2011; Segers & Tillema, 2011). It's believed that teachers' practices could be altered if changes are made to their beliefs (Barnes et al., 2015). However, there are a few existing studies that looked at both concurrently (Barnes et al., 2015; Ndalichako, 2015; Segers & Tillema, 2011).

James and Pedder (2006) investigated the relationship between teachers' perceptions of their classroom assessment practices and their assessment values. Five hundred and fifty eight teachers from infant through to secondary across 32 schools in England completed a questionnaire that was designed by the researchers. Teachers were to provide two kinds of response to the thirty items on the questionnaire. The first response was on assessment practices (never true, rarely true, often true, mostly true), while the second response asked about the importance they placed on the practice in creating opportunities for students to learn (not all important, limited important or important / crucial). Findings revealed that majority of the teachers placed high values on formative practices such as discussing with students ways to improve (93%), helping student to plan next step in learning (83%), allowing students to think about how they learn (95%), peer assessment (72%) and helping students to learn independently (92%), while a small percent placed value on summative practices such as norm references (20%), assessment consisting of primarily marks and grades (35%) and planning lessons based on the curriculum rather than previous assessment (51%).

James and Pedder (2006) classified the items into three key dimensions of classroom assessment practices; making learning explicit, promoting learning autonomy and performance orientation. Making learning explicit and promoting learning autonomy were associated with formative practices, while performance orientation was associated with summative practices. Using factor and cluster analysis, they discovered inconsistencies between practices and values. It was basically concluded that teachers used assessment practices that they do not value as important to students' learning. For example, 93% stated that they highly valued discussing with students ways to improve, but only 54% stated that they practiced such in their teaching. This accounted for a 39% gap between value and practice. Similar findings were found in relation to summative practices. For example, while only 20% indicated that they value norm referencing practices and 35% indicated that they value using marks and grades, a greater percentage (29% and 45% respectively) indicated that they used such practices. Thus, despite placing a high value on assessments that improve the quality of students' learning, low scores were recorded on teachers' actual assessment practices. James and Pedder (2006) attributed these discrepancies between values and practices to curriculum obligation and testing regime that existed in England.

Using a version of James and Pedder's (2006) questionnaire, Chew and Lee (2013) also conducted a study to investigate facilitators' beliefs and practices of classroom assessment. 148 facilitators from a polytechnic in Singapore completed an online survey. They also found inconsistencies between beliefs and practices. For instance, although teachers indicated that they did not believe in performance orientation practices such as using grades and marks primarily and using the curriculum to plan future lesson rather than previous assessment, they indicated using such practices. Contrary to James and Pedder (2006), Chew and Lee (2013) found that Singapore teachers did not strongly believe in formative practices, but still used such practices in their class. Overall, teachers' assessment practices scored higher than their beliefs.

Azis (2015) also found differences between teachers' beliefs and practices. Using a mixed method approach, 107 English teachers completed the TCoA inventory and those that scored high for improvement conceptions were interviewed in South Sulawesi, Indonesia. Teachers strongly advocated for formative practices that help students develop their learning, but they also favoured summative practices such as grading to show achievement. Azis (2015) concluded that the pattern of the teachers' responses from the TCoA survey was not consistent with their perceptions and practices. Teachers indicated that they favoured assessments that are used for improving teaching and learning, but their responses about their assessment practices revealed they more favoured accountability conceptions.

Contrary to inconsistencies findings, other studies have revealed that teachers' assessment beliefs match up with their assessment practices. Brown et al. (2009) found that teachers' conception predicted their practices. Improvement conceptions regressed on diagnostics and teaching improvement practices suggesting a relationship between improvement conceptions and formative practices. In the same way accountability conceptions regressed on school quality and examination practices and irrelevancy conceptions predicted irrelevant practices. Similar findings were found in Calveric (2010) who did a study to discovery third and fourth grade teachers' assessment beliefs and the importance they placed on different assessment practices. She found statistical relationship between student accountability conceptions and performance quiz, major exams and publisher test. Moderate relationship was also found between school accountability and major exams and publisher test. Improvement conceptions revealed the weakest relationship with major exams. All conceptions showed consistent findings with practices. Adam and Hsu (1998) also found consistent findings between conceptions and practices. They explored mathematics teachers' conceptions about assessment and their actual assessment practices using a cross-sectional survey. They found that the least important conception related to essay items, while the most important conception related to using observation. However, conception about students' performance was given the second highest importance. Assessment practices corresponded with assessment conceptions in the fact that essay item was given the lowest usage and teachers' observation was given the highest. Adams and Hsu (1998) concluded that despite using a variety of tools, teachers' conception of assessment is consistent with their assessment practices.

These findings reveal discrepancies between teachers' assessment conceptions and their assessment practices. Teachers generally believe that assessment should be conducted to improve teaching and learning, but they also believe in using assessment for accountability measures. Thus, findings show teachers valuing assessments that are used to improve teaching and learning, yet, using summative traditional practices. James and Pedder (2006) pointed out that a reason for the contradiction could be the testing regime in one’s country. They attributed the gap they found between teachers’ beliefs and practices to the testing context in England at the time that required teachers to push students to pass examination; while still promoting performance assessment. Azis (2015) also attributed the differences between Indonesia teachers’ beliefs and practices to the existing assessment system in Indonesia that focuses more on external examinations. A further reason for the contradiction as put forth by Azis (2015) could be that teachers are confused about assessment conceptions.

3.12 TEACHERS’ ASSESSMENT CONCEPTIONS AND PRACTICES IN BELIZE

Classroom assessment practices in Belize are guided by internal policies of individual school. While the Education Act of Belize dictates the use of strategies to improve learning and instruction, scant empirical evidence is available on assessment practices used by teachers and there is a non-existence of empirical evidence on teachers’ assessment conceptions. The only empirical study found during the search of literature was Armstrong (2011) who examined the impact of teachers’ training and use of assessment data among primary school teachers in the Belize district. She discovered no differences between teachers’ assessment practices and their use of assessment data, their level of training, gender, age and teaching experience. Her findings suggested that teacher with an associate degree in primary education had the highest understanding of assessment, highest usage of positive assessment practices and the highest usage of assessment data compared to teachers with a Master’s and Bachelor’s in education. This scarcity of empirical studies on teachers’ assessment practices and conceptions points to the need for research studies in the area.

3.13 SUMMARY

The main goal of education is to enhance students' learning with hope that it will increase students' achievement. Formative assessment is considered as the main tool to use to enhance students' learning. Thus, strategies such as providing students with detailed and effective feedback, allowing students to contribute to instruction planning and using assessment information to plan instruction have been shown to enhance learning. Such strategies are developed best through alternative practices such as feedback, peer assessment, self-assessment, performance assessment, authentic assessment, questioning and observations. Contrary to such recommendations, studies have revealed that teachers' assessment practices are still dominantly traditionally summative. This means that they still frequently use practices that cater to transmission styles of teaching. Such practices include multiple-choice, true and false and short response items. Teachers' conception of assessment has been seen as a main avenue to understand the practices used by teachers, but their conceptions varied as their practices. Moreover, teachers' conceptions about assessment do not necessarily predict their assessment practices.

Studying assessment may be problematic since the field is wide and encompasses many other areas; moreover, teachers' practices and conceptions may be the result of many unknown and known factors. It is in fact the varied and inconsistent usages and beliefs of assessment that makes the study of secondary teachers' assessment conceptions and practices in Belize extremely important. The common theme throughout this literature is that teachers' assessment conceptions and beliefs are based upon an individual country's context. A country's cultural, societal and educational contexts have significant impact on teachers' conceptions and practices. The fact that there is scant empirical evidence available on teachers' assessment practices and practically none on their conceptions in Belize further extent the significance of this study. Thus, this study main's focus is on providing empirical evidences on teachers' assessment conceptions and practices and to explore whether teachers' practices are related to their conceptions. The subsequent chapter outlines the methodology and procedures used to gather those empirical evidences.

CHAPTER FOUR

4.0 METHODOLOGY

4.1 INTRODUCTION

According to Neuman (2006), positivist approach to research utilizes a survey to gather quantitative data. This approach focuses on measurements and uses mathematical statistics to analyse the data. "Positivism is in accordance with empiricism view that knowledge stems from human experiences. It has an atomistic, oncological view of the world as comprising discrete, observable elements and events that interact in an observable, determined and regular manner" (Collins, 2010, p. 38). This research is concerned with studying teachers' assessment practices and beliefs from a quantitative methodology through the use of a survey. This chapter describes the research design, population, sampling, instrument, data collection and data analysis.

4.2 RESEARCH DESIGN

A quantitative research collects data in the form of number (Neuman, 2006). According to Creswell (2012), a quantitative study is best used when the research problem is based on trend. This allows the researcher to study the overall results of the respondents and how those results vary among people. This study will be examining secondary teachers' assessment practices and conceptions to discovery trends and variations. A quantitative study will allow for the examination of teachers' practices and conceptions and how those vary among gender, subject, years of experiences and assessment training. Further, a quantitative study may also be used to explain how one variable relates to another variable (Creswell, 2012). This study will also be looking at how teachers' assessment conceptions relate to their assessment practices.

4.2.1 Survey Research

Survey research is utilized in a study to learn about people's beliefs or opinions (Babbie, 2015; Creswell, 2012; Neuman, 2006). Neuman (2006) states that survey is “appropriate for research questions about self-reported beliefs and behaviour” (p. 273). It is a popular procedure in quantitative studies because it lends itself to statistical analysis and it is feasible. Surveys are able to gain information about large amount of people in a short time, it is able to ask many things at one time to measure many variables and it is cheap and easy to administer (Creswell, 2012; Neuman, 2006). Further, survey research allows for association among variables and could be used to relate results from data to results from past research studies (Creswell, 2012).

Survey research may be used for both descriptive and explanatory purposes (Babbie, 2015; Neuman, 2006). The purpose of a descriptive research is to describe people’s behaviour and gather their perceptions, opinions and beliefs about a current issue in education (Creswell, 2012; Lodico, Spaulding & Voegtle, 2006), while explanatory’s purpose is to explain why events occur and to build upon a theory (Babbie, 2015; Creswell, 2012). In this sense, an explanatory research is considered correlation when the objective is to explain the association between two variables and the extent to which two variables co-vary (Creswell, 2012; Lodico et al., 2006).

Survey is commonly used in studies on teachers’ assessment practices and beliefs. All of the studies reviewed on teachers’ assessment practices used a survey to gain an understanding of teachers’ practices. Similarly, a review of research studies on teachers’ conceptions of assessment revealed that four out of the six studies were quantitative (Azis, 2012). The researcher further stated that “survey research reflects the appropriate choice in investigating teachers’ conceptions of assessment” (Azis, 2012, p. 47). Thus, this research used a cross-sectional survey design to study teachers’ assessment conceptions and practices. Data was collected at one point in time to measure teachers’ current assessment practices and their current conceptions of assessment. This design also allowed the researcher to see the association between practices and conceptions.

4.3 SAMPLING

4.3.1 Setting

This study was conducted in government-funded high schools in the district of Belize. According to the MOE's (2015a) *Abstract of Education Statistics*, there are 56 high schools in the country of Belize of which 23 are owned and managed by different denominations, 12 are private / specialized, 4 are owned and managed by different communities and 17 are owned and managed by the government. The Belize district has the highest amount of government-funded high schools (7), follow by Stann Creek (3), Cayo (2), Orange Walk (2), Toledo (2) and lastly Corozal (1).

4.3.2 Population and sample

A population is a group of people that share similar characteristics (Creswell, 2012). Conclusions are made about the population based on the findings of a study (Babbie, 2015). The population being studied is all secondary teachers teaching in government-funded high schools on the mainland of the Belize District. The total estimated teaching force for this population was 191 (SIB, 2013). The target population was Mathematics, English Language, Social Studies and Science teachers in the six government funded high schools. Eighty nine teachers made up the target population and all teachers were invited to participate in the study. Out of the eighty nine teachers, fifty-eight completed the survey. This equated to a response rate of 65%.

Demographic data showed that 36% of the respondents were males and 64% were females. Thirty three percent of the teachers taught English Language, 27% taught Mathematics, 18% taught Science and 22% taught Social Studies. A total of 16% of the teachers have been teaching for less than three years, 33% between four to ten years, 42% between eleven to twenty years and 9 % more than twenty years. In addition, 78% have taken a university course in classroom assessment while 79% have received some in-service or workshop training in assessment, testing or measurement. In total, 19% of the respondents held a Master Degree, 55% a Bachelor Degree, 5% held a diploma, 5% a certificate and 16% chose other.

4.4 INSTRUMENT

A self-administered questionnaire was used in the study (see Appendix A). This type of questionnaire is completed by the respondent and is deemed very effective with a high response rate (Babbie, 2015; Neuman, 2006). In addition, results from a questionnaire could be summarized in charts, tables or graphs and analysed statistically (Neuman, 2006). The questionnaire used in the study had three parts and contained fifty six close-ended items. Part I asked the teachers about their beliefs and understanding of assessment, part II asked about teachers' assessment practices and their usage of assessment information and part III of the survey asked for demographic information.

Part I of the survey was adopted, with Dr. Gavin Brown's Permission (see Appendix G), from Brown's (2006) inventory of Teachers Conception of Assessment (TCoA) III Abridged version. The inventory comprised of nine factors, each made up of three items, and conceptualized four conceptions of assessment; assessment improves learning, assessment holds school accountable, assessment holds student accountable and assessment is irrelevant. The improvement conception comprised of four factors; assessment describes students' learning, assessment is valid, assessment improves students' learning and assessment improves teaching. These four factors made up twelve items in the questionnaire namely questions 3, 4, 5, 6, 12, 13, 14, 15, 21, 22, 23, and 24. According to Brown (2006), these items are grounded on the fact that assessment guides teachers and students on how to improve learning because it provides valid and accurate description of learning. On the other hand, the irrelevancy conception is comprised of three factors that are based on the fact that assessment is bad for student, ignored by teachers and is inaccurate (Brown, 2006). The three factors are assessment is unfair, assessment is ignored, and assessment is inaccurate which made up nine items in the questionnaire namely questions 7, 8, 9, 16, 17, 18, 25, 26, and 27. Only one factor made up of three items comprised the school accountability (1, 10, 19) and student accountability conceptions (2, 11, 20).

Teachers used a positive packed rating scale from the following range: strongly disagree, mostly disagree, slightly agree, moderately agree, mostly agree and strongly agree. Brown (2006) chose such a rating scale because it is expected that participants would be positively inclined to the various constructs; thus, a positive packed scale would show more variance within the positive range. The use of a positive packed scale was also recommended in Brown and Michaelides (2011) who used a balanced scale with teachers from Cyprus, but discovered that few teachers chose the “disagree somewhat” option. Thus, Brown and Michaelides (2011) recommended that future studies use the positive packed rating scale approach.

Part II of the survey (see Appendix A) has two sections that sought to find out the frequency of teachers using various assessment practices. The first section asked teachers to rate the frequency of various practices that were divided into two dimensions; traditional assessment and alternative assessment. Five items (30, 33, 40, 41, 35) made up the traditional dimension and included practices such as self-made test with select-type items and major exams, while ten items (28, 29, 31, 32, 34, 36, 37, 38, 39, 42) made up the alternative dimension and included practices such as individual projects, questioning and performance assessments. The second section of part II asked teachers about their usage of assessment information. Two dimensions were created to measure how teachers used assessment information. The formative dimension comprised of five items (43, 45, 46, 48, 50) and included usages such as to plan teaching and to help student address problems in learning. The summative dimension had three items (44, 47, 49) and included items such as for students’ promotion and graduation and to assign grades. Both sections of part II asked teachers to rate the frequency in which they used the various assessment practices using a six-point Likert scale. The scale ranged from not at all, very rarely, rarely, occasionally, frequently to very frequently.

The last part of the survey asked teachers about demographic information. Teachers were asked about their gender, subject taught, teaching experience, classroom assessment training, and educational qualification.

4.4.1 Reliability and Validity

Reliability refers to the measure of consistence, while validity means that the instrument is accurately measuring what it supposed to. The following steps where taken to ensure reliability and validity of the instrument used in the research. The abridged version of the Teachers' Conceptions of Assessment (TCoA) inventory was validated by Brown (2006). Exploratory Factor Analysis (EFA) confirmed that the items loading on each of the nine factors were sufficient (Brown, 2006). Confirmatory Factor Analysis (CFA) was also done which showed good fit characteristics (X_{311}^2 = 841.02; RMSEA = 0.57; TLI = .87) and similar interfactor correlation values and directions were reported. An independent confirmatory study was also done with a sample of Queensland teachers. A total of 1398 primary and high school teachers completed the CoA-III abridged version. The CFA showed a good fit (X_{933}^2 = 3283.56; RMSEA = .042; TLI = .81). These confirmatory analysis statistics allowed Brown (2006) to conclude that the "inventory is an efficient and valid measure of teachers' conception of assessment, and so it can be used within programs for teachers' professional development and research on assessment" (p. 170). Moreover, several studies from different geographical and educational contexts have used TCoA-III Abridged version to study teachers' assessment conceptions. These studies include Azis (2015), Brown et al. (2009) Brown et al. (2010), Brown et al. (2011), Brown and Michalis (2011) and Segers and Tillema (2011).

Items for the section on teachers' assessment practices were gathered from various studies on teachers' assessment practices including Zhang and Burry-Stock (2003), James and Pedder (2006), McMillan (2001) and Suah and Ong (2012). Content validity, how well the instrument measured the construct, was checked by two university professors in Fiji who are considered experts in the field of classroom assessment. The lectures reviewed the items to ensure clarity and checked to see that the items correctly covered the specific dimensions of assessment practices. Face validity was also established during this process. Thereafter, the necessary changes were made based on their recommendations.

The assessment practice section of the questionnaire had items based on teachers' assessment practices and their usage of assessment information. To test for reliability, a reliability coefficient test was using Cronbach's Alphas. This test measured how the accuracy of the instrument. An alpha score of 0.80 indicates a good internal consistency. The Cronbach's Alpha for this section was (α = .82).

4.5 DATA COLLECTION PROCEDURES AND CONSENT PROCESS

Permission and consent were sought from key players to conduct the study. Approval and consent were obtained from USP's Research Ethics Committee who approved the research intent and proposal of the study (Appendix B). A permission letter was sent to the General Manager of Government Schools who recommended that permission be sought from the principals within the six schools (see Appendix C). Letters were then sent to the respective principals to administer the questionnaire to the teachers in the first week of May (see Appendix D). Each principal was provided with an overview of the study and was informed about the research assistant. By the third week of May, all the principals have given consent to visit the schools to collect data from the respective teachers.

Due to the distance between the researcher and the sample population, a research assistant was recruited to distribute and collect the questionnaires. The research assistant was a current secondary school teacher who was chosen because of her competency and experience in conducting research. An interview was conducted via Skype between the researcher and research assistant to discuss the purpose of the study, the instrument and the process of distributing and collecting the questionnaire. A protocol was also developed that provided the assistant with a step by step process that was to be followed (see Appendix E).

After consent was given by the principals, the research assistant visited each school to distribute and collect the questionnaire. These visits occurred between the last two weeks in May and the first two weeks in June. All visits took place after regular class sessions had ended to provide teachers with sufficient time to complete the questionnaire and to minimize disruption. An overview of the study was discussed with teachers before soliciting consent to participate in the study. Each teacher was provided with a consent letter (see Appendix F) that described the purpose of the study, highlighted the benefits and explained expectations of teachers. Those who agreed to participate were given the questionnaire to complete in the presence of the research assistant who immediately collected the completed questionnaires. This process was repeated across the six secondary schools. All completed questionnaires were then mailed to the researcher.

4.6 ANALYSIS OF DATA

Data collected were coded and placed in the Statistical Package for the Social Science (SPSS) system. Variables for both assessment beliefs and assessment practices used an ordinal scale and were coded as follows: assessment beliefs (strongly disagree -1, mostly disagree - 2, slightly agree - 3, moderately agree - 4, mostly agree - 5 and strongly agree - 6) and assessment practices (not at all - 1, very rarely - 2, rarely - 3, occasionally - 4, frequently - 5 to very frequently - 6). The accuracy of the codes were checked by cleaning the data. Cleaning data simply means that the researcher looked for responses or answer categories that have errors (Neuman, 2006). This was done by checking all the variables for inaccurate codes.

The following illustrates how data were analysed for each research question.

What assessment practices are used by teachers in government funded high schools in Belize?

Two types of analysis were done to examine the frequency of teachers' assessment practices. Descriptive statistics were used to describe teachers' practices. Means, standard deviation, frequencies and percent were used to analyse the different teaching practices. The means is the average and is computed as the sum of all observed outcomes from the sample, the frequency is the rate at which something occurs and standard deviation is the measure of dispersion in a frequency distribution. The second analysis computed the scores for two new constructs; formative and summative assessments and then used descriptive statistics to describe the frequency of each construct.
What are secondary teachers' conceptions (beliefs) of assessment?

Descriptive statistics were also used to describe teachers' beliefs. Means, standard deviation, frequencies and percent were used to analyse the different conceptions. Pearson's Correlation was also used to determine the association among the four conceptions. This measure the strength of a linear association between two variables.

What are the differences among assessment beliefs and practices based on gender, teaching experience, measurement training and subject taught?

Analysis of Variance (ANOVA) and Independent Sampling *t* Test were used to analyse the differences among demographics information (independent variables) such as gender, teaching experience, measurement training and subject taught based on assessment practices and assessment beliefs (dependent variables). Both tests are used to determine whether there are any statistically significance between the means of independent groups. The difference is that ANOVA compares the means among more than two groups, while *t* Test compares only two groups.

How do teachers' beliefs of assessment relate to their assessment practices?

Pearson's Correlation Test was used to analyse the relationship between teachers' assessment beliefs and their practices.

4.7 SUMMARY

This chapter describes the research design, population, sampling, instrument, data collection and data analysis utilized in designing and conducting this study. The chapter discusses the purposes of a quantitative study and benefits of using a questionnaire to gather information on teacher's assessment practices and conceptions. A quantitative study allows the researcher to examine the overall responses of the participants and to see how their responses vary. It also allows the researcher to examine the association between variables. A questionnaire was utilized because it is able to gain information from a large group of people in a short time, it is cheap and it is easy to administer. Data gathered from the questionnaire was analysed using SPSS. The following chapter provides the findings of the study.

CHAPTER FIVE

5.0 FINDINGS OF THE STUDY

5.1 INTRODUCTION

The purpose of the study was to examine teachers' assessment conceptions and practices in Belize. It attempted to determine the types and frequency of assessment strategies used by teachers, how teachers use the information gathered from those assessments and their conceptions about assessment. The study made comparisons between teachers' assessment practices and conceptions and teacher characteristics' such as training, subject taught, years of experience and gender. The study also examined the relationship between teachers' conceptions of assessment and their assessment practices.

This chapter presents the findings of the study. The data were collected using a quantitative approach with the major purpose being descriptive and explanatory. The chapter is divided into five parts. The first part looks at the demographics of the respondents, followed by the second part that presents the findings on teachers' assessment practices. The third part presents the findings on teachers' assessment conceptions, the fourth part presents the results of teachers' assessment conceptions and practices in relation to gender, training, teaching experience and subject taught. The last part shows the results for the relationship between conceptions and practices.

5.2 RESPONE RATE:

The rate of response varied among the six schools as could be seen from Table 5.1. The largest amount of respondents (25%) came from school A and the second largest (13%) came from school B. As can be seen from Table 5.1, there was a 0% rate of return from school F. This however, did not affect the data since the sample population was the entire population of 89 teachers. Out of the 89 teacher population, 58 teachers completed the questionnaire which equalled to response rate of 65%.

Table 5.1: *Response Rate by Schools*

School	Completed Surveys	Percent Completion
School A	22	25%
School B	12	13%
School C	5	6%
School D	11	12%
School E	8	9%
School F	0	0%

5.3 DEMOGRAPHICS

A total of 58 secondary in-service teachers from government funded high schools in the Belize district made up the respondents who completed the survey instrument. This section will present the demographic information of those respondents based on sex, subject taught, teaching experience, measurement / assessment training and educational qualification.

Out of the 58 respondents, 37 were females and 21 were males (Table 5.2). Majority of teachers (18) taught mathematics, 15 taught English, 12 taught social studies and 10 taught science (Table 5.3). Three teachers did not provide information regarding their subject area; thus, only 55 responses yielded subject taught results. The respondents varied in their teaching experience with 42% teaching for 11-20 years, 33% teaching for 4-10 years, 16% teaching less than 3 years and 9% teaching for more than 20 years (Figure 5.1).

Table: 5.2. *Frequency and Percentage for Gender*

Gender	Frequency	Percent
Female	37	63.8
Male	21	36.2
Total	58	100.0

Table: 5.3. *Frequency and Percentage for Subject Taught*

Subject Taught	Frequency	Percent
Math	15	27.3
English	18	32.7
Social Studies	12	21.8
Science	10	18.2

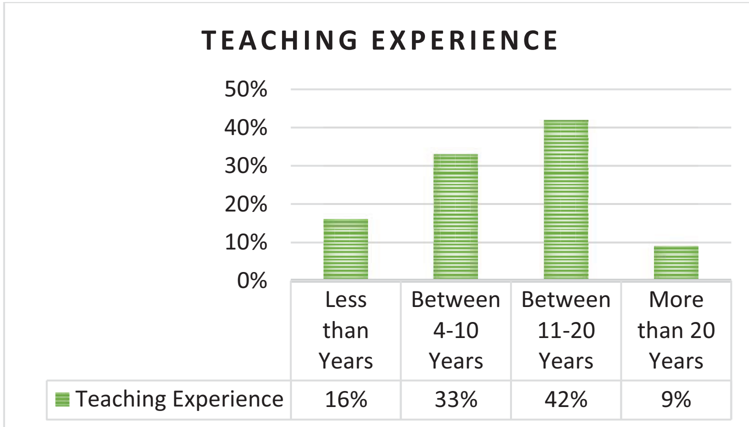

Figure: 5.1. Bar Chart Showing Percentage for Teaching Experience

Table 5.4 shows that majority of the respondents (45) had taken a university course in classroom assessment, while only 13 of them did not take such course. Similarity results were seen regarding those who received in-service or workshop training in assessment with the majority (46) indicated that they received such training and 12 indicated that they did not. In total, 19% of the respondents held a Master Degree, 55% a Bachelor Degree, 5% held a diploma, 5% a certificate and 16% chose other (see Figure 5.2).

Table: 5.4. *Frequency and Percentage for Classroom Assessment Training*

Training	Yes	No
University Course	45 (78%)	13 (22%)
In-service / Workshop	46 (79%)	12 (21%)

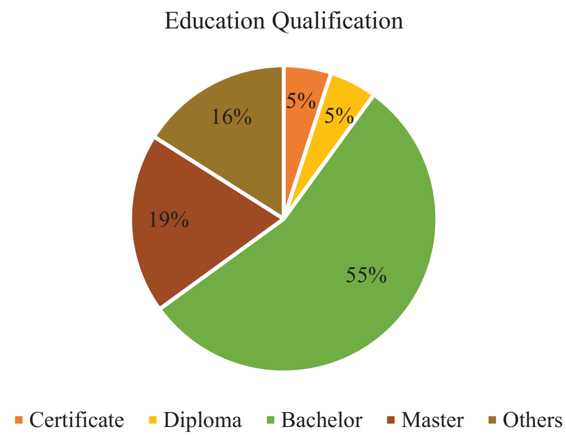

Figure: 5.2. Pie Chart Showing Education Qualifications

5.4 ASSESSMENT PRACTICES USED BY TEACHERS

Items related to assessment practices were divided into two parts; the first part asked about different types of assessment practices and the second part asked about the usage of assessment information. Two types of analysis were done to examine the frequency of assessment practices used by the respondents. The first analysis presented the frequency and percentage of participants' responses to different assessment practices and uses of assessment information. The second analysis computed the scores for different items to create two additional constructs; formative assessment practices and traditional assessment practices. Descriptive statistics, means and standard deviations were then computed to ascertain the frequency of each construct.

5.4.1 Teachers' Assessment Practices

Table 5.5 shows the frequency and percentage of teachers' assessment practices. Combination of percentage for *frequently* and *very frequently* responses revealed that majority of the respondents (79%) indicated that they frequently used questioning, 62% indicated that they frequently used self-made test with select type items such as multiple choices, true and false and matching, 58% frequently used observation, 57% frequently used group project and 50% frequently used major exams. Overall, majority of the respondents indicated that they *occasionally* used all of the following practices: self-assessment (34%), individual project (52%), major exams (43%), oral presentations (41%), assessments provided by publishers (43%), portfolio (36%), peer assessment (43%), authentic assessment (34%), self-made test with supply items such as short answers (36%), essay (38%) and performance assessment (40%).

A combination of percentage for *very rarely* and *rarely* responses showed that assessment practices such as self-assessment (30%), portfolio (39%), authentic assessment (31%), assessment provided by publishers (26%) and performance assessment (20%) were rarely used by the respondents. Interestingly, 12% indicated that they did not use self-assessment. None of the respondents chose *not at all* for individual projects, major exams, group projects, oral presentation and self-made test with supply-type items suggesting that the practices are commonly used among the teachers.

Table: 5.5. *Frequency and Percentage for Assessment Practices*

Assessment Practices	**Not at All**		**Very Rarely**		**Rarely**		**Occasionally**		**Frequently**		**Very Frequently**	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Self-Assessment	7	12%	9	16%	8	14%	20	34%	12	21%	2	3%
Individual Project	0	0%	1	2%	3	5%	30	52%	20	34%	4	7%
Major Exams	0	0%	2	3%	2	3%	25	43%	14	24%	15	26%
Observation	4	7%	1	2%	3	5%	16	28%	19	33%	14	25%
Group Project	0	0%	0	0%	2	3%	23	40%	25	43%	8	14%
Test with Select-Type Items	3	5%	3	5%	5	9%	11	19%	23	40%	13	22%
Oral Presentations	0	0%	2	3%	7	12%	24	41%	19	33%	6	10%
Provided by Publishers	1	2%	3	5%	12	21%	25	43%	12	21%	5	9%
Portfolio	4	7%	7	12%	15	27%	20	36%	8	14%	2	4%
Peer Assessment	4	7%	5	9%	5	9%	25	43%	16	28%	3	5%
Questioning	2	3%	0	0%	5	9%	5	9%	25	43%	21	36%
Authentic Assessment	2	3%	2	3%	16	28%	20	34%	13	22%	5	9%
Test with Supply-type Items	0	0%	2	3%	7	12%	21	36%	20	34%	8	14%
Essay	5	9%	4	7%	4	7%	21	38%	16	29%	6	11%
Performance Assessment	3	5%	1	2%	10	18%	22	40%	12	22%	7	13%

5.4.2 Usage of Assessment Information

Frequently and *very frequently* yielded high rate for items in this section (see Table 5.6). For instance, 53% indicated that they used assessment information for promotion and graduation *very frequently*, 53% indicated that they *frequently* used it to provide descriptive feedback and 50% indicated that they *frequently* used it to plan teaching. A joint review of usages rated *frequently* and *very frequently* revealed that 84% of the respondents used assessment information for promotion and graduation, 84% used it to assign grades, 83% used it to plan teaching, 82% to provide descriptive feedback, 77% used it to help students address problems in learning, 76% used it to rank students and 57% used it to make decisions regarding individual students. Unlike assessment practices, few respondents chose *occasionally* in this section; 32% indicated that they *occasionally* used assessment without providing a grade and 31% indicated that they *occasionally* made decisions about individual students.

Surprisingly, none of the respondents chose *rarely*, *very rarely* or *not at all* for using assessment to plan teaching or help students address problems in their learning, suggesting that these practices are occasionally or frequently used. Interesting, 23% chose *not at all* and 20% indicated that they *very rarely* used assessment without providing grades. A combination of *not at all*, *very rarely* and *rarely* percentages yielded 48% of respondents who indicated that they rarely used or never used assessments without providing a grade.

Table: 5.6. *Frequency and Percentage for Usage of Assessment Information*

Usage of Assessment Information	**Not at All**		**Very Rarely**		**Rarely**		**Occasionally**		**Frequently**		**Very Frequently**	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Make Decision about Individual Student	2	3%	0	0%	5	9%	18	31%	23	40%	10	17%
Promotion and Graduation	0	0%	0	0%	4	7%	5	9%	18	31%	31	53%
Plan Teaching	0	0%	0	0%	0	0%	10	17%	29	50%	19	33%
Help Students Address Problems in Learning	0	0%	0	0%	0	0%	13	23%	27	47%	17	30%
Assign Grade	0	0%	1	2%	1	2%	7	12%	23	40%	25	44%
Provide Descriptive Feedback	1	2%	1	2%	0	0%	8	14%	31	53%	17	29%

Rank Student	1	2%	0	0%	4	7%	9	16%	18	31%	26	45%
Assessment without Providing Grade	13	23%	11	20%	3	5%	18	32%	7	12%	4	7%

5.4.3 Comparison between Traditional and Formative Assessment Practices

As mentioned earlier, two additional constructs were created to examine the frequency of usages based on traditional assessment practices and formative assessment practices. The formative assessment variable was created based on items relating to formative practices identified in the literature. Eleven items made up the construct: self-assessment, observations, oral presentations, portfolio, peer assessment, questioning, performance assessment, use assessment to make decision about individual student, use assessment to plan teaching, use assessment to help student address problems in learning and use assessment without providing grades. All items used a 6-points Likert scale ranging from *not at all* to *very frequently*. To test the inter-relatedness of the 11 items, a Cronbach's alpha test was done which yielded an alpha of .803. This indicated a high level of interrelatedness of the items. Descriptive statistics calculated yielded mean = 4.24 and SD = .677 (see Figure 5.3). Figure 5.3 also illustrates the variance of the items that made up formative assessment variable. Majority of the items fell within the 4 – 5 range suggesting that majority of the teachers indicated that they occasionally or frequently used formative assessment practices.

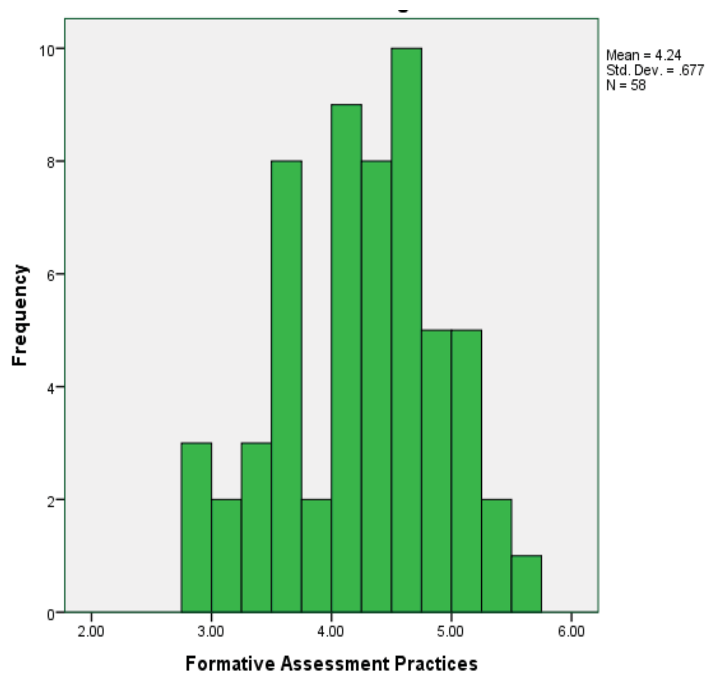

Figure: 5.3. Bar Graph showing the Variance in Response Based on Formative Construct

The traditional assessment variable was created based on items relating to traditional practices identified in the literature. Seven items made up the variable: major exams, self-made test with select-type items such as multiple choices, true / false and matching, self-made test with supply-type items such as short answers, assessment provided by publishers, use assessment to assign grades, use assessment to rank students and use assessment for students' promotion and graduation. Items also used a 6-points Likert scale ranging from *not at all* to *very frequently*. Unlike formative assessment variable, the inter-relatedness of the 7 items that made up the traditional variable was poor with a Cronbach's alpha of .604 which indicated a low level of interrelatedness. The mean was 4.7 with a SD of .611 (see Figure 5.4).

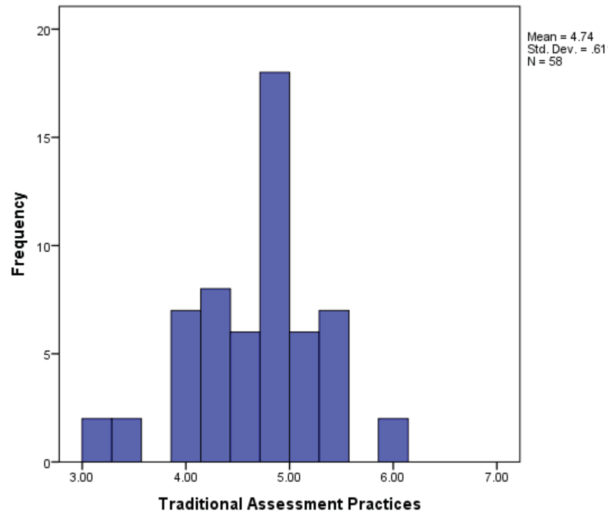

Figure: 5.4. Bar Chart Showing the Variance in Response Based on Traditional Construct

5.5 TEACHERS' CONCEPTIONS ABOUT ASSESSMENT

Descriptive statistics were used to determine the means, standard deviation, frequency and percentage of teachers' conceptions about assessment. The instrument for teachers' conceptions of assessment was adopted from Brown's (2006) CoA-III Abridged who divided the 27 items into four conceptions of assessment namely; assessment makes school accountable, assessment makes student accountable, assessment improves education and assessment is irrelevant. Table 5.7. outlines how the items were grouped into the four conceptions.

Table: 5.7. *Four Conceptions of Assessment*

Four Conceptions of Assessments	Items
School Accountability	Provides information on how well schools are doing
	An accurate indicator of a schools' quality
	Good way to evaluate a school
Student Accountability	Places students into categories
	Assigning a grade or level to student work
	Determines if students meet qualification standards
Improvement Accountability	Determine how much students have learned from teaching
	Establishes what students have learned
	Measures students higher order thinking skills
	Provides feedback to students about their performance
	Feedbacks to students their learning needs
	Helps students improve their learning
	Integrated with teaching practices
	Modifies outgoing teaching of students
	Allows different students to get different instruction
	Results are trustworthy
	Results are consistent
	Results can be depended on

Irrelevancy	Forces teachers to teach against their beliefs Unfair to students Interferes with teaching Make little use of the results Results are filed and ignored Little impact on teaching Results should be treated cautiously given measurement error Error and imprecision in all assessment Imprecise process

Information was adapted from Table 1: Conceptions of Assessment-III: First- and Second-Order Factors, Statements, and Loadings for Two Jurisdictions. Adapted from "Teachers' Conceptions of Assessment: Validation of an Abridged Version," by Gavin T. L. Brown, 2006, *Psychological Reports*, 99(1), 166-170.

Descriptive statistics were then run to determine the mean and standard deviation for each conception (Table 5.8). The improvement and student conceptions yielded the highest means of 4.3 and 4.3 respectively suggesting that majority of the respondents moderately agreed with both conceptions. The standard deviation is used to indicate the extent to which the values are clustered around the mean or spread away from it. Improvement conception yielded a SD of .76 while student accountability yielded SD of .91. Both SD are low indicating that majority of responses clustered around the mean. The lowest mean scores were yielded for school accountability (mean = 3.8) and irrelevancy (mean = 2.9). The mean score for school accountability indicates that majority of the respondents slightly agreed with the conception while the mean score for irrelevancy suggests that majority of the respondents mostly disagreed with it. While irrelevancy yielded a SD of .72 suggesting that majority of the responses clustered around the mean, school accountability yielded the highest SD (1.0) of all the conceptions indicating some variance within responses. Figure 5.5 shows that 50% of the respondents fell within the range of *slightly agree – moderately agree*, while 25% fell within the range of mostly agree – strongly and 25% fell within *mostly disagree* - *slightly disagree.*

Table: 5.8. *Means and Standard Deviations for Assessment Conceptions*

Assessment Conceptions	*N*	Mean	Std. Deviation
Improvement	58	4.3311	.76331
Student Accountability	58	4.3046	.91401
School Accountability	58	3.8420	1.05521
Irrelevancy	58	2.9862	.72195

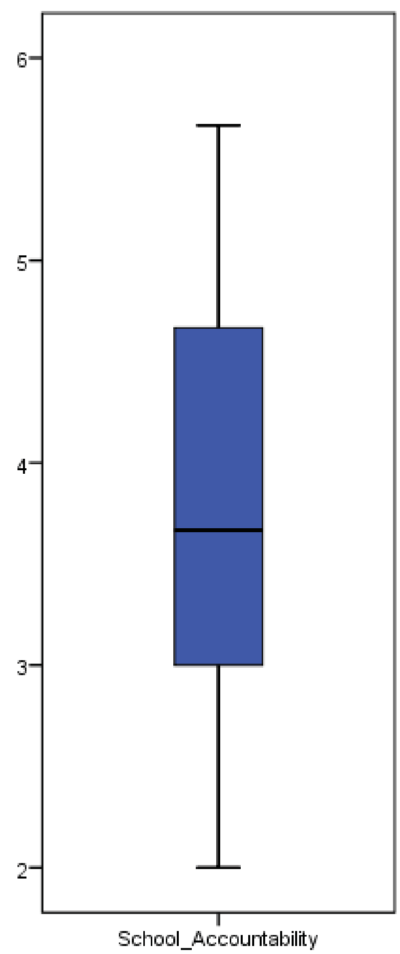

Figure: 5.5. Box Plot Showing Variance of Responses for School Accountability

5.5.1 School Accountability Conception

Percent was computed for items under each conception. Table 5.9 illustrates the three items that made up school accountability conception. 34% of the teachers mostly agreed that assessments provide information on how well schools are doing, 21% slightly agreed, 19% strongly agreed, 17% moderately agreed, 9% mostly disagreed and 0% strongly disagreed. Collectively, mostly agree and strongly agree accounted for 53% of the responses, slightly agree and moderately agree accounted for 38% and the disagreed responses accounted for 9% suggesting that majority of the respondents believe that assessment is a way to determine how well schools are doing. Results for assessment being an accurate indicator of school's quality were slightly different with mostly and strongly agree accounting for only 36% of the responses, slightly and moderately agree accounted for 33% and strongly and mostly disagree accounted for 31% of the responses indicating that some teachers do not really believe that assessment is a good indicator of school quality. The last item within school accountability conception yielded varying responses with 9% strongly disagreed that assessment is a good way to evaluate schools, 14% mostly disagreed, 22% slightly agreed, 21% moderately agreed, 17% mostly agreed and 17% strongly agreed. Overall, the majority agreed with school assessment conception.

Table: 5.9. *Frequency and Percent for Items within School Accountability Conception*

School Accountability	**Strongly Disagree**		**Mostly Disagree**		**Slightly Agree**		**Moderately Agree**		**Mostly Agree**		**Strongly Agree**	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
How well schools are doing	0	0%	5	9%	12	21%	10	17%	20	34%	11	19%
Indicator of School's Quality	7	12%	11	19%	12	21%	7	12%	18	32%	2	4%
Good Way to Evaluate a School	5	9%	8	14%	13	22%	12	21%	10	17%	10	17%

5.5.2 Student Accountability Conception

Three items also made up the student accountability conception (Table 5.10). Majority of the respondents either mostly agreed (31%) or strongly agreed (29%) with using assessment to place students into categories. Majority moderately agreed (28%) or mostly agreed (31%) with assessment being assigning a grade to student's work. The statement 'assessment determines if students meet qualification standards' had the highest amount of respondents choosing agreement responses.

Despite high percentage of agreement, there was a small percentage of disagreement with the items in this conception.

Table: 5.10. *Frequency and Percent for Items within Student Accountability Conception*

Student Accountable	**Strongly Disagree**		**Mostly Disagree**		**Slightly Agree**		**Moderately Agree**		**Mostly Agree**		**Strongly Agree**	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Places students into categories	1	2%	6	10%	5	9%	11	19%	18	31%	17	29%
Assigning Grade to Students Work	3	5%	6	10%	10	17%	16	28%	18	31%	5	9%
Determine if Students Meet Qualification Standards	0	0%	3	5%	12	21%	9	16%	26	45%	8	14%

5.5.3 Improvement Conception

Twelve items made up the improvement conception (Table 5.11). Moderately agree, mostly agree and slightly agree yielded the highest percentage of responses. Respondents strongly agreed with the following statements: assessment being used to determine how much students have learned (40%), to provide feedback to student (58%) and integrating assessment with teaching practices (55%). Respondents also mostly agreed with using assessment to modify ongoing teaching (43%), to measure higher order thinking (35%) and to improve student's learning (30%). Thirty nine percent moderately agreed with assessment results being trustworthy, 34% moderately agreed that assessment establishes what students have learned, and 36% moderately agreed that assessment feeds back to students their learning needs. A small percentage of the respondents disagreed with items in the improvement conception. For instance, a combination of strongly and mostly disagree responses yielded 39% disagreeing with assessment results are consistent and 22% disagreeing that assessment results can be depended on.

Table: 5.11. *Frequency and Percent for Items within Improvement Conception*

Improvement	Strongly Disagree		Mostly Disagree		Slightly Agree		Moderately Agree		Mostly Agree		Strongly Agree	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Determine how much students have learned	0	0%	2	3%	8	14%	8	14%	17	29%	23	40%
Provides feedback to students	1	2%	2	4%	0	0%	7	12%	14	25%	33	58%
Integrated with teaching practices	0	0%	2	4%	2	4%	4	7%	17	30%	31	55%
Results Are Trustworthy	2	4%	3	5%	10	18%	22	39%	17	30%	3	5%
Establishes What Students Have Learned	0	0%	6	10%	6	10%	20	34%	17	29%	9	16%
Feed Backs to Student Their Learning Needs	0	0%	3	5%	6	10%	21	36%	16	28%	12	21%
Modifies Ongoing Teaching of Students	0	0%	4	7%	4	7%	13	22%	25	43%	12	21%
Results are Consistent	4	7%	18	32%	13	23%	15	26%	5	9%	2	4%
Measures Higher Order Thinking	0	0%	2	4%	16	28%	12	21%	20	35%	7	12%
Improve Students' Learning	0	0%	4	7%	11	19%	12	21%	17	30%	13	23%
Different Instruction	0	0%	9	16%	18	31%	16	28%	8	14%	7	12%
Can be Depended on	4	8%	7	14%	15	29%	10	20%	13	25%	2	4%

5.5.4 Irrelevancy Conception

Nine items made up the irrelevance conception which received the highest percentage of disagreement (see Table 5.12). A combination of the two disagreement responses showed a high percent of disagreement. Majority of the respondents disagreed that teachers make little use of assessment results (42%), that it is unfair to student (67%), that it is filed and ignored (55%), that it interferes with teaching (70%), that it has little impact on teaching (74%) and that it is an imprecise process (55%). A combination of the four agreement responses and the two disagreement responses paints another story. For instance, while 41% of the respondents disagreed that assessment forces teachers to teach against their beliefs, 59% agreed. Similar results were yielded for making little use of assessment results with 42% disagreement and 58% agreement. Furthermore, 92% agreed that assessment results should be treated cautiously because of measurement error and

90% agreed that teachers should take into account the error and imprecision in all assessment.

Table: 5.12. *Frequency and Percent for Items within Irrelevancy Conception*

Irrelevancy	Strongly Disagree		Mostly Disagree		Slightly Agree		Moderately Agree		Mostly Agree		Strongly Agree	
	N	%	N	%	N	%	N	%	N	%	N	%
Forces Teachers to Teach against Beliefs	9	16%	14	25%	18	32%	8	14%	5	9%	2	4%
Make Little Use of Results	8	14%	16	28%	11	19%	12	21%	10	18%	0	0%
Measurement Error in Assessment	0	0%	5	9%	5	9%	19	33%	16	28%	13	22%
Unfair to Students	18	32%	20	35%	11	19%	5	9%	0	0%	3	5%
Filed and Ignored	13	22%	19	33%	14	24%	7	12%	2	3%	3	5%
Account of Error and Imprecision	2	3%	4	7%	7	12%	9	16%	18	31%	18	31%
Interferes with Teaching	15	27%	24	43%	8	14%	5	9%	2	4%	2	4%
Little Impact on Teaching	18	32%	24	42%	8	14%	4	7%	2	4%	1	2%
Imprecise Process	13	23%	18	32%	11	19%	10	18%	3	5%	2	4%

5.5.5 Relationship between Conceptions

Pearson's correlation analysis was done to determine the association among the four conceptions of assessment (Table 5.13). Strong positive associations were seen between improvement and student accountability (r = 0.61, p = .000), improvement and school accountability (r = 0.66, p = .000) and student accountability and school accountability (r = 0.58, p = .000). A negative weak association (r = -0.13) was yielded for improvement and irrelevancy indicating that as the respondent agreed with improvement, they disagreed with irrelevancy. A weak association (r = 0.12) was also found between school accountability and irrelevancy.

Table: 5.13. *Correlations Analysis of Assessment Conceptions*

Assessment Conceptions		Improvement	Student Accountability	School Accountability	Irrelevancy
Improvement	r	1	.610**	.664**	-.134
	p		.000	.000	.318
	N	58	58	58	58
Student Accountability	r	.610**	1	.589**	.148
	p	.000		.000	.267
	N	58	58	58	58
School Accountability	r	.664**	.589**	1	.120
	p	.000	.000		.370
	N	58	58	58	58
Irrelevancy	r	-.134	.148	.120	1
	p	.318	.267	.370	
	N	58	58	58	58

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

5.6 DIFFERENCES BETWEEN CONCEPTIONS AND PRACTICES AND TEACHERS' CHARACTERISTICS

Analysis of Variance (ANOVA) and Independent Sampling *t*-test were computed to determine if differences existed between teachers' characteristics and assessment conceptions and practices. Composite mean scores were computed using descriptive statistics for assessment conceptions and practices. ANOVA and *t*-test were then conducted to determine if the differences in the mean scores of the different groups of teachers are incidental or statistical significant.

5.6.1 Gender

Independent sampling *t*-test was conducted between gender and assessment practices to determine if there was any significant differences between the two. Table 5.14 shows the mean score for sex in relation to the different assessment practices. As could be seen from the table, there was no huge difference in mean scores between male and female teachers in relation to the different assessment practices. Only three items; questioning, authentic assessment and essay, yielded slight differences among mean scores. *t*-test yielded *p*-values of over .05 for all assessment practices indicating no significant differences between male and females when it comes to assessment practices (refer to Table H1 in Appendix H).

Table: 5.14. *Differences between Male and Female for Assessment Practices*

Assessment Practices	Sex	*N*	Mean	Std. Deviation
Self-Assessment	Female	37	3.3784	1.40141
	Male	21	3.6190	1.39557
Individual Project	Female	37	4.3784	.79412
	Male	21	4.4286	.74642
Major Exams	Female	37	4.6216	.98182
	Male	21	4.7143	1.10195
Observation	Female	37	4.4595	1.50175
	Male	20	4.6500	1.03999
Group Project	Female	37	4.5946	.76229
	Male	21	4.8095	.74960
Test with Select-type Items	Female	37	4.5676	1.21428
	Male	21	4.3810	1.59613
Oral Presentations	Female	37	4.2973	.96796
	Male	21	4.4286	.92582
Provided by Publishers	Female	37	4.1351	1.05836
	Male	21	3.8095	1.07792
Portfolio	Female	35	3.5143	1.31443
	Male	21	3.4286	1.02817
Peer Assessment	Female	37	3.9189	1.18740
	Male	21	3.9048	1.37495
Questioning	Female	37	5.1892	1.04981
	Male	21	4.5714	1.28730
Authentic Assessment	Female	37	4.0270	1.11770
	Male	21	3.8095	1.20909
Test with Supply-type Items	Female	37	4.5676	1.04191
	Male	21	4.1905	.87287
Essay	Female	37	3.9730	1.42374
	Male	19	4.1053	1.32894
Performance Assessment	Female	36	4.0833	1.38099
	Male	19	4.1053	.87526

No statistical significant differences were yielded between the mean scores of male and female teachers' usage of assessment information (refer to Table H2 in Appendix H). Table 5.15 shows that majority of mean scores between male and female teachers were similar with the exception of assigning grade, providing descriptive feedback, ranking students and assessment without grade which yielded slight differences.

Table: 5.15. *Differences between Male and Female Teachers for Usage of Assessment Information*

Usage of Assessment Information	Sex	*N*	Mean	Std. Deviation
Make Decision about Individual Student	Female	37	4.5405	1.04335
	Male	21	4.5714	1.20712
Promotion and Graduation	Female	37	5.4324	.76524
	Male	21	5.0952	1.09109
Plan Teaching	Female	37	5.1622	.68773
	Male	21	5.1429	.72703
Help Students Address Problems in Learning	Female	36	5.0556	.71492
	Male	21	5.0952	.76842
Assign Grade	Female	36	5.3889	.59894
	Male	21	4.9524	1.16087
Provide Descriptive Feedback	Female	37	5.1351	.82199
	Male	21	4.8571	1.10841
Rank Student	Female	37	5.1622	.92837
	Male	21	4.9524	1.32198

Assessment without Providing Grade	Female	36	3.2500	1.59239
	Male	20	2.9000	1.68273

Table 5.16. shows the mean scores for the four conceptions of assessment between male and female teachers. Improvement, student accountability and school accountability yielded similar mean scores between male and female, while irrelevancy yielded a slightly different mean score between the two with male respondents yielding a slightly higher mean score (3.1) than female respondents (2.8). Independent sample *t*-test was also conducted to determine the differences in mean scores between male and female respondents in relation to assessment conceptions. The *p*-values for the four conceptions were larger than .05 which indicated that there was no significant differences in mean score for female and male respondents in their agreement or disagreement regarding the conceptions (refer to Table H3 in Appendix H).

Table: 5.16. *Differences between Male and Female Teachers for Assessment Conceptions*

Conceptions	Sex	*N*	Mean	Std. Deviation
Improvement	Female	37	4.4183	.73657
	Male	21	4.1775	.80320
Student Accountability	Female	37	4.4324	.81219
	Male	21	4.0794	1.05359
School Accountability	Female	37	3.8919	1.03951
	Male	21	3.7540	1.10255
Irrelevancy	Female	37	2.8973	.76285
	Male	21	3.1429	.63053

5.6.2 Subject Taught

ANOVA was used to examine the mean differences among the four subject areas. ANOVA is similar to *t*-test except that it is able to compare more than three groups while *t*-test is only able to compare between two groups. ANOVA tells which group, if any, is statistically significantly different than the other groups. Table 5.17 shows the mean scores for each assessment practice in relation to the four subject areas. Self-assessment, individual project, group project, oral presentation, portfolio and self-made test with supply-type items yielded similar mean scores across the four subjects. Science teachers recorded the highest mean scores for major exams, observation and self-made test with select-type items, while mathematics teachers yielded the lowest mean scores for essay and performance assessment.

Table: 5.17. *Differences in among Subject Areas for Assessment Practices*

Assessment Practices		*N*	Mean	Std. Deviation
Self-Assessment	Math	15	3.6667	1.63299
	English	18	3.5556	1.19913
	S.Studies	12	3.5000	1.16775
	Science	10	3.2000	1.68655
Individual Project	Math	15	4.2667	.59362
	English	18	4.3333	.97014
	S.Studies	12	4.5833	.51493
	Science	10	4.5000	.97183
Major Exams	Math	15	4.4667	1.06010
	English	18	4.6667	1.13759
	S.Studies	12	4.4167	.66856
	Science	10	5.2000	1.13529
Observation	Math	14	4.3571	1.64584
	English	18	4.3333	1.64496
	S.Studies	12	4.7500	.86603
	Science	10	5.0000	.81650
Group Project	Math	15	4.6000	.91026
	English	18	4.5556	.78382
	S.Studies	12	4.8333	.57735
	Science	10	4.9000	.73786
Test with Select-type Items	Math	15	4.2667	1.48645
	English	18	4.0556	1.39209
	S.Studies	12	4.5000	1.31426
	Science	10	5.6000	.69921
Oral Presentations	Math	15	4.0000	.92582
	English	18	4.5556	.78382
	S.Studies	12	4.4167	.90034
	Science	10	4.5000	1.26930
Provided by Publishers	Math	15	4.2000	1.01419
	English	18	4.2778	1.22741
	S.Studies	12	3.5833	.66856
	Science	10	4.2000	.78881
Portfolio	Math	15	3.5333	1.06010
	English	17	3.8824	1.36393
	S.Studies	11	3.2727	.64667
	Science	10	3.4000	1.26491
Peer Assessment	Math	15	3.8000	1.52128
	English	18	4.3333	.97014
	S.Studies	12	3.6667	.88763
	Science	10	4.1000	1.28668
Questioning	Math	15	4.8667	.99043
	English	18	5.2778	1.17851
	S.Studies	12	4.6667	1.55700
	Science	10	5.1000	.87560
Authentic Assessment	Math	15	3.6667	1.23443
	English	18	4.2222	.80845
	S.Studies	12	3.7500	1.28806
	Science	10	4.2000	1.39841
Test with Supply-type Items	Math	15	4.1333	1.06010
	English	18	4.6111	.97853
	S.Studies	12	4.1667	1.11464
	Science	10	4.9000	.73786
Essay	Math	15	2.9333	1.57963
	English	18	4.4444	1.04162
	S.Studies	11	4.5455	.93420
	Science	10	4.6000	1.07497
Performance Assessment	Math	15	3.6000	1.40408
	English	18	4.2778	1.27443
	S.Studies	11	4.2727	.90453
	Science	10	4.4000	1.07497

ANOVA test revealed significant difference for the usage of self-made test with select-type items (p = 0.031) and essay (p = .001) (refer to Table H4 in Appendix H).

Sidak's Post Hoc test was done to determine the location of the differences (refer to Table H5 in Appendix H). A significant difference was seen between English and science teachers (p = .02) for self-made test with select-type items and between math and all other subjects English (p = .005), science (p = .008), social studies (p = .009) for essay.

Mean scores for using assessment information to make decisions about individual students and for promotion and graduation were similar among the four subject areas (see Table 5.18). English teachers yielded the lowest mean score (4.9) for using assessment to plan teaching, while social studies teachers yielded the lowest mean score for using assessment information to help student address problems in learning (4.5), assigning grade (4.6) and ranking students (4.7). English (4.9) and science (4.9) yielded the lowest mean scores for providing descriptive feedback and math (2.6) and science (2.8) teachers yielded the lowest means for using assessment without providing grade.

Table: 5.18. *Differences among Subject Areas for Usage of Assessment Information*

Usages of Assessment Information		*N*	Mean	Std. Deviation
Make Decision about Individual Student	Math	15	4.8667	1.35576
	English	18	4.4444	1.14903
	Social Studies	12	4.5000	.90453
	Science	10	4.5000	.84984
Promotion and Graduation	Math	15	5.5333	.83381
	English	18	5.3889	.84984
	Social Studies	12	5.0000	.95346
	Science	10	5.4000	.84327
Plan Teaching	Math	15	5.6000	.50709
	English	18	4.9444	.72536
	Social Studies	12	5.0000	.60302
	Science	10	5.2000	.78881
Help Students Address Problems in Learning	Math	14	5.6429	.49725
	English	18	5.0556	.63914
	Social Studies	12	4.5833	.66856
	Science	10	5.0000	.81650
Assign Grade	Math	15	5.6000	.63246
	English	18	5.4444	.61570
	Social Studies	11	4.6364	1.28629
	Science	10	5.1000	.73786
	Total	54	5.2593	.87276
Provide Descriptive Feedback	Math	15	5.3333	.61721
	English	18	4.9444	.99836
	Social Studies	12	5.2500	.62158
	Science	10	4.9000	.73786
Rank Student	Math	15	5.4667	.91548
	English	18	5.2778	.82644
	Social Studies	12	4.7500	1.05529
	Science	10	5.0000	1.05409
Assessment without Providing Grade	Math	15	2.6667	1.75933
	English	17	3.7059	1.61108
	Social Studies	11	2.8182	1.66242
	Science	10	3.2000	1.31656

ANOVA test revealed significant differences for plan teaching (*p* = .035), help students address problem in learning (*p* = .002) and assigning grade (*p* = .024) (refer to Table H6 in Appendix H). Sidak's post hoc test revealed statistical significant differences between math and English teachers (*p* = .038) for plan teaching, between social studies and math teachers (*p* = .001) for helping student address problems in learning and between math and social studies teachers (*p* = .027) for assigning grades (refer to Table H7 in Appendix H).

In relation to assessment conceptions, science teachers yielded the lowest mean scores for all four conceptions; improvement (3.7), student accountability (3.8), school accountability (3.1), and irrelevancy (2.8) (Table 5.19). English teachers yielded the highest mean score for school accountability (4.0) and student accountability (4.5), while social studies teachers yielded the highest for improvement (4.4) and irrelevancy (3.1). ANOVA test revealed no significant difference among subject areas in relation to assessment conceptions (refer to Table H8 in Appendix H).

Table: 5.19. *Differences among Subject Areas for Assessment Conceptions*

Assessment Conceptions		*N*	Mean	Std. Deviation
Improvement	Math	15	4.4237	.67904
	English	18	4.4373	.83039
	Social Studies	12	4.4729	.54255
	Science	10	3.7970	.92168
Student Accountability	Math	15	4.3333	1.02353
	English	18	4.5000	.76055
	Social Studies	12	4.3889	.83887
	Science	10	3.8667	1.12437
School Accountability	Math	15	3.8444	1.03791
	English	18	4.0926	1.04666
	Social Studies	12	3.9444	1.03312
	Science	10	3.1667	.97183
Irrelevancy	Math	15	2.9800	.82393
	English	18	2.9278	.61911
	Social Studies	12	3.1917	.78678
	Science	10	2.8500	.80863

5.6.3 Teaching Experience

Table 5.20 shows the mean scores among the different groups of teaching experience. Self-made test with supply-type items, self-assessment, individual project, group project, oral presentation, portfolio, questioning and authentic assessment are all used on a similar frequency regardless of teaching experience. Teachers with less than 3 years of experience yielded a slightly lower mean score for observation (3.8) and self-made test with select-type items (3.7) than teachers in the other subgroups. Teachers with more than 20 years of experience yielded the lowest mean score (3.4) for assessment provided by publisher and teachers with 4 – 10 years of experience yielded the highest mean for usage of major exams (5). ANOVA test revealed a significant difference with the usage of performance assessment (p = .031) (refer to Table H9 in Appendix H); however, Sidak's post hoc test revealed a high p-value of .053 between teachers with less than three years and teachers with 4 – 10 years of experience (refer to Table H10 in Appendix H).

Table: 5.20. *Differences among Teaching Experience for Assessment Practices*

Assessment Practices	Teaching Experience	*N*	Mean	Std. Deviation
Self-Assessment	Less than 3 years	9	3.3333	1.41421
	Between 4 - 10 Years	19	3.8421	1.42451
	Between 11 - 20 Years	24	3.2917	1.45898
	More than 20 Years	5	3.0000	1.00000
Individual Project	Less than 3 years	9	4.5556	.88192
	Between 4 - 10 Years	19	4.3684	.76089
	Between 11 - 20 Years	24	4.4167	.65386
	More than 20 Years	5	4.0000	1.22474
Major Exams	Less than 3 years	9	4.6667	1.00000
	Between 4 - 10 Years	19	5.0000	1.10554
	Between 11 - 20 Years	24	4.4167	.82970
	More than 20 Years	5	4.4000	1.51658
	Total	57	4.6491	1.02628
Observation	Less than 3 years	9	3.8889	2.26078
	Between 4 - 10 Years	19	4.7895	1.03166
	Between 11 - 20 Years	24	4.5000	1.21584
	More than 20 Years	4	4.7500	.95743
Group Project	Less than 3 years	9	4.5556	.88192
	Between 4 - 10 Years	19	4.8421	.83421
	Between 11 - 20 Years	24	4.6250	.71094
	More than 20 Years	5	4.6000	.54772
Test with Select-type Items	Less than 3 years	9	3.7778	1.30171
	Between 4 - 10 Years	19	4.8421	1.42451
	Between 11 - 20 Years	24	4.4167	1.41165
	More than 20 Years	5	4.8000	.44721
Oral Presentations	Less than 3 years	9	4.1111	1.05409
	Between 4 - 10 Years	19	4.3158	1.15723
	Between 11 - 20 Years	24	4.5417	.77903
	More than 20 Years	5	4.0000	.70711
Provided by Publishers	Less than 3 years	9	4.7778	1.09291
	Between 4 - 10 Years	19	3.8421	.95819
	Between 11 - 20 Years	24	4.0000	1.02151
	More than 20 Years	5	3.4000	1.34164
Portfolio	Less than 3 years	8	3.7500	1.66905

	Between 4 - 10 Years	19	3.4211	1.34643
	Between 11 - 20 Years	23	3.3913	1.07615
	More than 20 Years	5	3.6000	.54772
Peer Assessment	Less than 3 years	9	4.5556	.88192
	Between 4 - 10 Years	19	4.0526	1.35293
	Between 11 - 20 Years	24	3.6667	1.23945
	More than 20 Years	5	3.4000	1.34164
Questioning	Less than 3 years	9	4.6667	1.41421
	Between 4 - 10 Years	19	4.9474	1.07877
	Between 11 - 20 Years	24	4.9583	1.23285
	More than 20 Years	5	5.6000	.89443
Authentic Assessment	Less than 3 years	9	4.0000	1.41421
	Between 4 - 10 Years	19	4.0526	1.22355
	Between 11 - 20 Years	24	3.7083	1.08264
	More than 20 Years	5	4.4000	.54772
Test with Supply-type Items	Less than 3 years	9	4.6667	.86603
	Between 4 - 10 Years	19	4.5263	.84119
	Between 11 - 20 Years	24	4.2083	1.17877
	More than 20 Years	5	4.8000	.83666
Essay	Less than 3 years	9	4.0000	1.41421
	Between 4 - 10 Years	19	3.7895	1.47494
	Between 11 - 20 Years	23	4.2609	1.32175
	More than 20 Years	5	3.8000	1.48324
Performance Assessment	Less than 3 years	9	3.1111	1.36423
	Between 4 - 10 Years	18	4.3889	1.03690
	Between 11 - 20 Years	23	4.0870	1.20276
	More than 20 Years	5	4.8000	.83666

The mean scores for using assessment to make decisions about individual student, for promotion and graduation and for assigning grade were similar among the different groups (Table 5.21). The results revealed that teachers with more than 20 years of experience used assessment to plan teaching (4.8) and provide descriptive feedback (4.4) slightly less than the other subgroups who all yielded mean scores above (5). The mean score for using assessment to help students address problems in their learning was lower for teachers with more than 11 years of experience than the mean score for teachers with less than 11 years of experience. Differences were also seen between the mean scores for using assessment to rank students for teachers with less than 3 years (4.6) and teachers in the other groups whose mean scores were all above 5. ANOVA test revealed no significant differences with the usage of assessment information among the different groups (refer to Table H11 in Appendix H).

Table: 5.21. *Differences Teaching Experience for Usage of Assessment Information*

Usage of Assessment Information	Teaching Experience	*N*	Mean	Std. Deviation
Make Decision about Individual Student	Less than 3 years	9	4.2222	1.48137
	Between 4 - 10 Years	19	4.6842	1.29326
	Between 11 - 20 Years	24	4.6667	.81650
	More than 20 Years	5	4.2000	.83666
Promotion and Graduation	Less than 3 years	9	5.0000	1.00000
	Between 4 - 10 Years	19	5.2105	1.08418
	Between 11 - 20 Years	24	5.4583	.77903
	More than 20 Years	5	5.6000	.54772
Plan Teaching	Less than 3 years	9	5.0000	.70711

	Between 4 - 10 Years	19	5.3684	.68399
	Between 11 - 20 Years	24	5.1667	.70196
	More than 20 Years	5	4.8000	.44721
Help Students Address Problems in Learning	Less than 3 years	9	5.1111	.60093
	Between 4 - 10 Years	19	5.3158	.74927
	Between 11 - 20 Years	23	4.9565	.76742
	More than 20 Years	5	4.8000	.44721
Assign Grade	Less than 3 years	9	5.2222	.83333
	Between 4 - 10 Years	19	5.3158	.88523
	Between 11 - 20 Years	23	5.1739	.98406
	More than 20 Years	5	5.2000	.44721
Provide Descriptive Feedback	Less than 3 years	9	5.0000	.86603
	Between 4 - 10 Years	19	5.0526	1.22355
	Between 11 - 20 Years	24	5.1667	.56466
	More than 20 Years	5	4.4000	1.34164
Rank Student	Less than 3 years	9	4.6667	1.41421
	Between 4 - 10 Years	19	5.0526	1.26814
	Between 11 - 20 Years	24	5.2500	.84699
	More than 20 Years	5	5.2000	.83666
Assessment without Providing Grade	Less than 3 years	9	3.1111	2.14735
	Between 4 - 10 Years	18	2.3333	1.41421
	Between 11 - 20 Years	23	3.5652	1.37597
	More than 20 Years	5	3.8000	1.78885

Table 5.22 shows the mean scores among the different groups of teaching experience for the four assessment conceptions. Teachers with less than 3 years of experience scored the highest mean for improvement (4.8), student accountability (4.5) and school accountability (4.3), while teachers with more than 20 years of experience scored the highest mean for irrelevancy (3.3). Teachers with more than 20 years of experience had the lowest mean for improvement (3.8), while teachers with 4 – 10 years of experience had the lowest mean for student accountability (4.0). Teachers with more than four years of experience all had mean scores of 3.7 for school accountability except teachers with less than three years of experience who had mean score of 4.3. A similar trend was seen for irrelevancy whereby teachers with less than 20 years of experience yielded similar mean score of 2.9 while teachers with more than 20 years of experience yielded a mean score of 3.3. ANOVA revealed no significant difference among the different groups of teaching experience in relation to assessment conceptions (refer to Table H12 in Appendix H).

Table: 5.22. *Differences in Mean Scores among Teaching Experience for Assessment Conceptions*

Assessment Conceptions	Teaching Experience	*N*	Mean	Std. Deviation
Improvement	Less than 3 years	9	4.8005	.66873
	Between 4 - 10 Years	19	4.2408	.72984
	Between 11 - 20 Years	24	4.3378	.78257
	More than 20 Years	5	3.8970	.81090
Student Accountability	Less than 3 years	9	4.5185	.91456
	Between 4 - 10 Years	19	4.0000	1.03638
	Between 11 - 20 Years	24	4.4583	.84449
	More than 20 Years	5	4.4667	.69121
School Accountability	Less than 3 years	9	4.3704	1.20698
	Between 4 - 10 Years	19	3.7368	1.09758

		N	Mean	Std. Deviation
	Between 11 - 20 Years	24	3.7639	.97544
	More than 20 Years	5	3.7333	1.14018
Irrelevancy	Less than 3 years	9	2.9000	.82006
	Between 4 - 10 Years	19	2.9789	.58365
	Between 11 - 20 Years	24	2.9250	.78699
	More than 20 Years	5	3.3800	.84380

5.6.4 Assessment / Measurement Training

Respondents were asked to indicate whether they had taken a university course in classroom assessment or measurement training. Majority of the respondents (45) indicated that they took such a course, while 13 indicated that they did not. A *t*-test was conducted to examine the differences between the two groups of teachers in relation to assessment practices and conceptions. Table 5.23 shows the mean scores between the two groups for assessment practices. Teachers who had taken a university course yielded the highest mean scores for all assessment practices except major exams, assessment provided by publisher and self-made test with supply-type items. Significant differences were revealed via *t*-test for the usage of observation (p = .038) and performance assessment (p = .006) indicating that teachers who took a course in assessment used those two practices more frequently (refer to Table H13 in Appendix H).

Table: 5.23. *Differences in Assessment Training for Assessment Practices*

Assessment Practices	Classroom Assessment / Measurement Courses	*N*	Mean	Std. Deviation
Self-Assessment	Yes	45	3.6222	1.31924
	No	13	2.9231	1.55250
Individual Project	Yes	45	4.4222	.81153
	No	13	4.3077	.63043
Major Exams	Yes	45	4.6222	1.02888
	No	13	4.7692	1.01274
Observation	Yes	44	4.7273	1.24571
	No	13	3.8462	1.51911
Group Project	Yes	45	4.7111	.75745
	No	13	4.5385	.77625
Test with Select-type Items	Yes	45	4.5333	1.27208
	No	13	4.3846	1.66024
Oral Presentations	Yes	45	4.3778	.96032
	No	13	4.2308	.92681
Provided by Publishers	Yes	45	3.9556	.95240
	No	13	4.2308	1.42325
Portfolio	Yes	44	3.5227	1.22927
	No	12	3.3333	1.15470
Peer Assessment	Yes	45	3.9333	1.28629
	No	13	3.8462	1.14354
Questioning	Yes	45	5.0444	1.12726
	No	13	4.6923	1.31559
Authentic Assessment	Yes	45	4.0889	1.08339
	No	13	3.4615	1.26592

Test with Supply-type Items	Yes	45	4.3778	1.00654
	No	13	4.6154	.96077
Essay	Yes	44	4.1591	1.21890
	No	12	3.5000	1.83402
Performance Assessment	Yes	43	4.3256	1.12802
	No	12	3.2500	1.21543

Mean scores for usage of assessment information could be seen in Table 5.24. The two groups yielded similar mean scores for majority of the items including using assessment to make decision about individual student, for promotion and graduation, to plan teaching, to help student address problems in learning, to assign grade and to provide descriptive feedback. Teachers who took a university course in assessment yielded the highest mean scores for using assessment to rank students (5.1), and using assessment without providing grades (3.3). *t*-test revealed no significant differences between the two groups of teachers in relation to the usage of assessment information (refer to Table H14 in Appendix H).

Table: 5.24. *Differences in Assessment Training for Usage of Assessment Information*

	Classroom Assessment / Measurement Courses	*N*	Mean	Std. Deviation
Make Decision about Individual Student	Yes	45	4.6444	.98062
	No	13	4.2308	1.42325
Promotion and Graduation	Yes	45	5.4000	.83666
	No	13	5.0000	1.08012
Plan Teaching	Yes	45	5.1778	.68387
	No	13	5.0769	.75955
Help Students Address Problems in Learning	Yes	44	5.0682	.72810
	No	13	5.0769	.75955
Assign Grade	Yes	44	5.2273	.83146
	No	13	5.2308	1.01274
Provide Descriptive Feedback	Yes	45	5.0222	.78303
	No	13	5.0769	1.38212
Rank Student	Yes	45	5.1333	.86865
	No	13	4.9231	1.65638
Assessment without Providing Grade	Yes	43	3.3488	1.61664
	No	13	2.3846	1.44559

Table 5.25 provides the mean scores for the four conceptions based on assessment training course. Similar mean scores could be seen for both improvement, student accountability and irrelevance conceptions between the two groups of teachers. *t*-test revealed a significant difference (p = .008) for school accountability between the two groups (refer to Table H15 in Appendix H). Teachers who have not taken a university course in classroom assessment yield a higher mean score (4.5) than teachers who had taken a course (3.6).

Table: 5.25. *Differences in Assessment Training for Assessment Conceptions*

Assessment Conception	Classroom Assessment / Measurement Courses	*N*	Mean	Std. Deviation
Improvement	Yes	45	4.2584	.78561
	No	13	4.5828	.64511
Student Accountability	Yes	45	4.2519	.89374
	No	13	4.4872	.99643
School Accountability	Yes	45	3.6481	.98487
	No	13	4.5128	1.05071
Irrelevancy	Yes	45	3.0267	.72218
	No	13	2.8462	.73214

5.7 ASSOCIATION BETWEEN CONCEPTIONS AND PRACTICES

Pearson Correlation test was computed to examine the association between teachers' assessment conceptions and their assessment practices. Brown's (2006) TCoA Abridged III used a six point Likert scale ranging from strongly disagreed to strongly agreed, while the assessment practices instrument used a six point Likert scale ranging from not at all to very frequently.

Correlation was computed between all assessment practices and the four assessment conceptions (Table 5.26). Statistically significant relationship were seen between all four conceptions and specific assessment practices. Improvement conception yielded statistical significant relationship with the following assessment practices: assessment provided by publisher (r = .30, p = .021), using assessment information to plan teaching (r = .29, p = .014) and using assessment information to help student address problems in learning (r = 0.32, p = .016). School accountability produced statistically significant relationship with major exams (r = .25, p = .050) and assessment provided by the publisher (r = .37, p = .004). Student accountability also yielded significant relationship with using assessment information for promotion and graduation (r = .36, p = .005), using assessment information to rank student (r = .29, p = .023) and using assessment to assign grade (r = .30, p = .021). Statistical significant relationships were also detected between irrelevancy and authentic assessment (r = .30, p = .022) and using assessment without providing grade (r = .28, p = .035).

Table: 5.26. *Correlation Analysis between Conceptions and Practices*

Assessment Practices		Improvement	Student Accountability	School Accountability	Irrelevancy
Self-Assessment	*r*	.108	-.058	.057	.167
	p	.418	.664	.671	.210
	N	58	58	58	58
Individual Project	*r*	.035	.182	.068	.152
	p	.793	.170	.614	.255
	N	58	58	58	58
Major Exams	*r*	.228	.077	.259*	-.135
	p	.086	.565	.050	.311
	N	58	58	58	58
Observation	*r*	-.060	.116	-.121	.186
	p	.659	.392	.371	.167
	N	57	57	57	57
Group Project	*r*	.070	.138	.033	.117
	p	.603	.301	.807	.383
	N	58	58	58	58
Self-Made Test with Select Type Items	*r*	-.207	-.031	-.134	-.005
	p	.119	.819	.316	.968
	N	58	58	58	58
Oral Presentations	*r*	.057	-.022	.102	.107
	p	.669	.869	.445	.423
	N	58	58	58	58
Provided by Publishers	*r*	.302*	.246	.371**	-.057
	p	.021	.062	.004	.673
	N	58	58	58	58
Portfolio	*r*	.023	.024	.016	.019
	p	.865	.859	.908	.889
	N	56	56	56	56
Peer Assessment	*r*	-.002	-.002	.065	.094
	p	.988	.987	.628	.482
	N	58	58	58	58
Questioning	*r*	-.137	-.132	-.061	.043
	p	.304	.322	.647	.748
	N	58	58	58	58
Authentic Assessment	*r*	.014	.060	.039	.300*
	p	.919	.655	.771	.022
	N	58	58	58	58
Self-made Test with Supply-type Items	*r*	.095	.156	.150	.104
	p	.480	.243	.262	.438
	N	58	58	58	58
Essay	*r*	.212	.057	.161	.121
	p	.117	.676	.235	.376
	N	56	56	56	56
Performance Assessment	*r*	-.114	-.124	-.080	.240
	p	.407	.367	.560	.078
	N	55	55	55	55
Make Decision about Individual Student	*r*	.222	.191	.107	.083
	p	.094	.150	.423	.535
	N	58	58	58	58
Promotion and Graduation	*r*	.188	.366**	.145	.026
	p	.157	.005	.279	.849
	N	58	58	58	58
Plan Teaching	*r*	.293*	.320*	.070	-.038
	p	.026	.014	.603	.779
	N	58	58	58	58
Help Students Address Problems in Learning	*r*	.318*	.216	.164	-.089
	p	.016	.107	.223	.511
	N	57	57	57	57
Assign Grade	*r*	.023	.304*	.035	-.232
	p	.863	.021	.796	.083

	N	57	57	57	57
Provide Descriptive Feedback	r	.146	.240	.100	-.178
	p	.276	.069	.454	.180
	N	58	58	58	58
Rank Student	r	.115	.299*	.176	.053
	p	.392	.023	.186	.691
	N	58	58	58	58
Assessment without Providing Grade	r	-.096	-.075	-.072	.283*
	p	.479	.581	.598	.035
	N	56	56	56	56

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

*. Correlation is significant at the 0.05 level (2-tailed).

5.8. SUMMARY

This chapter presents the findings of the study on teacher's assessment conceptions and practices. Teachers indicated that they use an array of assessment practices including major exams, self-made test with select response, observation and group project. Teachers also indicated that they frequently used information from assessment to assign grade, to rank students, to plan teaching and to help student address problems in their learning. Practices such as authentic assessment, performance assessment, self-assessment and using assessment without providing a grade were rarely used by teachers. Assessment practices corroborates with assessment conceptions since majority of the teachers agreed with using assessment to make students accountable and to improve teaching and learning. Statistical differences were also yielded for specific practices and conceptions and teacher's demographic features.

CHAPTER SIX

6.0 DISCUSSION OF RESULTS

6.1 INTRODUCTION

Countries are moving away from quantity and now focusing on quality in order to develop competent individuals who are able to not only meet local standards, but are also able to compete within the global arena. The focus on quality education has seen many developed countries altering their assessment policies to enhance teaching and learning in the hope that students' achievements will increase. These countries look upon classroom assessment as a major tool to improve teaching and learning, since the use of specific assessment practices have been proven to strengthen and empower students to take charge of their own learning and achievement. Unfortunately, many developing countries are lagging behind when it comes to assessment. This is seen in the few empirical evidences existing on in-service teachers' assessment practices in developing countries. The lack of local empirical evidence means that policies on assessment are either based upon foreign evidence or based on anecdotal evidence. It is critical to not only have empirical evidence on the frequently used assessment practices among in-service teachers, but also to know teachers' beliefs about those practices so that the relevant policies are put in place and the relevant training are provided to teachers regarding assessment.

Changes in the classroom will only occur when teachers are empowered. In order to empower teachers we must be familiar with their practices and beliefs. Using a self-reporting questionnaire, in-service teachers within government high schools in the district of Belize rated their level of usage for specific assessment practices and also indicated their level of agreement or disagreement to a list of assessment conceptions. This chapter analyses and discusses the major findings from the study.

6.2 ASSESSMENT PRACTICES

The study found that teachers used a wide array of assessment practices. Majority of the assessment practices were used at a certain frequency. Practices such as questioning, observation, group project, major exams and self-made test with select-type responses such as multiple choice, true and false, fill in the blanks were given the highest percent of *frequently* responses suggesting that these practices are used on a high rate by teachers. Authentic assessment, assessment from publisher and performance assessment received the highest percent of *rarely* responses. Even though the overall response is indicating that teachers use a variety of practices, only five practices out of the fifteen listed practices did not receive any *not at all* responses. This suggests that individual project, major exams, group projects, oral presentations and self-made test with supply-type items are commonly used among all of the teachers since none indicated that they never used these practices.

With regards to usage of assessment information, majority of the teachers indicated that they frequently used information from assessment for promotion and graduation, to provide students with descriptive feedback, to plan teaching, to assign grade, to help students address problems in their learning, to rank students and to make decision regarding individual students. None of the teachers indicated that they never or rarely used information from assessment to provide descriptive feedback, to plan teaching or to help students address problems in learning suggesting that these practices are often used. Interesting; however, is that 23% indicated that they never and 25 % indicated that they rarely used assessments without providing grades.

These findings are in accordance with other studies done on teachers' classroom assessment practices which found similar findings. It has been repeatedly revealed that teachers seem to be fond of specific assessment practices. Bol et al. (1998), Gullickson (1985), Stiggins and Bridgeford (1985) and Marso and Pigge (1988) all reported practices such as self-made objective test, multiple choice, true and false, observation and major exams were highly used by majority of teachers. Fast-forward to thirty years and we find Frey and Schmitt (2010), Suan and Ong (2012) and Ogunkola and Clifford (2013) finding similar results. These studies have not only spanned over different time periods, but also spanned over different countries. This demonstrates that regardless of time and context, teachers have hold onto to these practices.

This study then corroborates those studies in revealing that teachers' assessment practices have been unchangeable over several years regardless of time or space. It is interesting that both developed and developing countries are on the same tangent when it comes to assessment practices. One would expect to find advance practices being used by teachers from more developed countries since they have far more opportunities, resources and technical experts than less developed countries. However, that does not seem to be the case since this set of ancient assessment practices has been consistent across time and space. These results verify what Shepard (2000) and others have long exposed; changes in teachers' practices need to go deeper than providing teachers with a workshop or a course in assessment practices.

Majority of the frequently used practices revealed in this study have not been acclaimed as the best practices for improving students' learning and achievement; yet they are frequently utilized by teachers. Practices that fall within the formative framework are considered the best assessment practices to improve learning and achievement. Contrary to these are practices that fall within the summative / traditional framework which are mainly used for accountability purposes and are not seen to improve learning since they are usually done at the end of a unit. To determine the frequency in which teachers in government high schools in Belize used these practices, a formative construct and a traditional construct were created. The formative construct yielded a high alpha score, while the alpha score for the traditional construct was low. This indicated that the items within the formative construct had high interrelationship among each other while the interrelationship among the items within the traditional construct was poor. Both constructs yielded low standard deviation indicating that majority of the responses were centred around the mean score. The traditional construct yielded a higher mean score than the formative construct suggesting that teachers use traditional practices more often than formative practices.

The actual count of responses paints a better picture of the variation in responses between the usage of traditional and formative practices. For instance, forty nine out of the fifty eight teachers indicated that they frequently used assessment information for promotion and graduation and to assign a grade, forty four indicated that they frequently used it to rank students and twenty nine indicated that they frequently used assessments such as major exams. Only one out of the above four practices had one teacher indicating that he or she never used assessment to rank students, four indicated that they rarely used it for promotion and graduation, two indicated that they rarely used it to assign a grade, four indicated that they rarely used it to rank students and only four indicating that they rarely used major exams. Contrary to the high usage of traditional practices is the lower usage of formative practices. Only eleven out of the fifty eight teachers indicated that they frequently used self-assessment and nineteen indicated they frequently used peer assessment. Compared to the traditional practices, a higher amount of teachers indicated that they never or rarely used assessment information to make decision about individual students and without proving a grade. These results indicate that while teachers do use certain formative assessment practices, they tend to prefer traditional practices.

As the common practices identified before, these findings on formative and summative assessment practices are also consistent with other findings in literature. McMillan et al. (2010) reported that despite finding that majority of teachers used formative assessment practices, only a small percentage of those teachers reported using specific practices such as using assessment to diagnose student weakness and to guide future instruction. Song and Koh (2010) discovered that the most frequently used formative practice was questioning. Other studies over the past thirty years have consistently revealed low usages of formative practices such as portfolio, feedback and self-assessment (Bol et al., 1998; Frey & Schmitt, 2010; Ongunkola & Clifford, 2013; Suan & Ong, 2012).

What is interesting about this study's findings on formative practices is that a high percentage of teachers indicated that they frequently used specific formative practices. Unlike McMillan et al. (2010), 48 out of the fifty eight teachers indicated that they use assessment information to plan teaching, while none indicated that they rarely or never used such practice. This high usage of specific formative assessment verses the low usages of other formative practices such as using assessment without providing grade makes one wonder if teachers are actually aware of the different purposes related to the different assessment practices. For instance, teachers indicating that they frequently use assessment to plan teaching may simply be saying that they plan future instruction solely on whether or not majority of students fail a particular test. It necessarily does not mean that teachers do systematic analysis on all assessment given on a particular topic to determine students' weak areas and then gear future instruction on those weaknesses. Similarly, the findings are indicating that teachers frequently used practices such as questioning and observation; however, the actual usage is unknown. What teachers believe to be observation and questioning might be different from what is considered effective observation and questioning. For instance, it may be common for teachers to observe that their students are on tasked; thus they indicated that they frequently use observation. In the same way, teachers may usually ask one or two questions during their lesson and thus indicated that they frequently use questioning. Thus, indicating that they frequently use observation and questioning does not necessarily indicate that they are aware of the purpose of those practices or that they use those practices effectively.

The idea that teachers may not be aware of the different purposes for different assessment practices are not new, but has always been a concern in literature. Studies have found that teachers use a consistent amount of practices across subjects and levels indicating that they fail to alter practices based on different purposes (Frey & Schmitt, 2010; Stiggins & Bridgeford, 1985; Zhang & Burry-Stock, 2003). Dayal and Lingam (2015) also found that despite associating assessment to formative function, majority of the teachers defined assessment in relation to a summative function.

Summative practices have been established as the most common and most frequently used assessment practices. Song and Koh (2010) found that teachers' summative assessment practices overshadowed their formative practices, Volante et al. (2010) discovered that teachers overused summative practices such as tests and quizzes.

The high usage of summative practices have been credited to the historical usage of assessment. For a long while assessment focused solely on the end result, measuring what students have learned. This belief is still common among teachers who see assessment as being outside of teaching and learning. This is why practices such as self-assessment, peer assessment and using assessment without providing grades are not seen as part of the teaching / learning sphere; thus, they are rarely used within the classroom.

6.2.1 Differences in Assessment Practices and Usage of Assessment Information Based on Teachers' Characteristics

Even though slight differences were seen in mean scores for assessment practices based on the different characteristics, few statistical differences were yielded. Statistical significant differences were not seen between male and female teachers in relation to assessment practices and usage of assessment information. This suggests that regardless of gender, teachers use similar assessment practices at similar frequency. With the exception of performance assessment, no other significant differences were seen among the different years of teaching experience for usage of assessment information. Teachers with more than four years of experience seem to use performance assessment at a higher frequency than those with less experience. Similar results were seen for measurement training. There was no significant differences between teachers who took a university course in measurement / assessment training and those who did not in relation to usage of assessment information, but differences were seen in the usage of assessment practices. The usage of observation and performance assessment were significantly differences between the two groups of teachers. Teachers who took a course in measurement / assessment training tend to use those practices more often than those that did not take a course.

There is no clear consensus on the differences between teachers' characteristics based on the usage of assessment practices. As Frey and Schmitt (2010), this study also found that more experienced teachers use performance assessment more often than less experienced teachers. However, Mertler (1998) found the opposite; novice teachers using alternative methods more than experienced teachers. Others have found no significant differences in the usage of assessment practices based on teaching experience.

Similar inconsistencies are seen with assessment practices and training. It is a widely held belief that trained teachers are better equipped to increase students' learning and achievement. One of the major education initiative in Belize currently is getting all primary and secondary school teachers trained. Zhang and Burry-Stock (2003) found that significant differences between teachers' training and assessment practices which is similar to the findings in this study regarding the use of observation and performance assessment. They even made the claim that training does impact the assessment practices that teachers use. Ogunkola and Clifford (2013); on the other hand, discovered that training does not play a role on the usage of assessment practices.

These inconsistencies may be due to methodology differences. Participants in the different studies not only came from different educational contexts, but the sampling size differed for each study. Sampling size has been shown to greatly impact statistical significance test. A larger sampling size is able to produce greater precision and statical power. Sampling size may also affect Type I and Type II errors (Babbie, 2015). Even though the entire population was surveyed in this study, the response rate was not as high as it could have been which would have provided for a larger sampling size for each group of characteristics. Nonetheless, exposing teachers to a university course in measure and assessment training can only strengthen their knowledge about assessment. In the same way, teachers who have been in the classroom for a longer time are better able to choose the most effective practice for a particular purpose. Thus, teacher training and teaching experience are thought to have an impact on teachers' assessment practices.

The claim has been made that assessment practices differ according to subject area. Many studies have found that mathematics and science teachers' assessment practices differ from English and science teachers (Duncan and Noonan, 2007; Frey & Schmitt, 2010; Ogunkola & Clifford, 2013; Ohlsen, 2007; Stiggins & Bridgeford, 1985). Mathematics and science teachers are thought to use objective assessment more often than other teachers. This study found statistical significant differences among the four subject areas for only two assessment practices. It was revealed that science teachers use test with select-type items such as multiple choice, true and false and fill in the blanks more often than English teachers. It was also revealed that mathematic teachers use essay less often than English Language, science and social studies teachers. Significant differences were also seen among the subjects based on the usage of assessment information. Differences were seen between mathematic and English for using assessment information to plan teaching, between mathematics and social studies for using it to address problems in student learning and between mathematics and social studies for using it to assign grade. In all three instances, mathematics teachers were found to use those practices more frequently than other subject teachers.

As explained by McMillan (2014), effective assessment practices could be utilized by all teachers regardless of the subject area. For instance, this study found significant differences with the usages of essay between math teachers and the other subject teachers. Even though essay is considered a traditional assessment practice, it is also considered effective in developing students' critical thinking skill and could also be used in mathematics lessons.

6.3 ASSESSMENT CONCEPTIONS

Despite countless research and years of appeals, summative practices still dominate teachers' classroom assessment practices. Examining the beliefs teachers hold regarding assessment may help to understand why this is the case. Knowing these beliefs may shed light on the choice of assessment practices teachers prefer. Studies have consistently found that teachers out rightly believe that assessment should be used to improve teaching and learning (Azis, 2015; Brown et al., 2009; Brown et al., 2010; Brown et al., 2011; Brown & Michaelides, 2011; Remesal, 2008; Segers & Tillema, 2011). This study corroborates those findings since majority of the teachers agreed with the improvement conception. However, majority of the teachers also agreed with student accountability conception. In fact, the mean scores were exactly the same for both improvement and student accountability conceptions. This is suggesting that while teachers believe that assessment should be used to improve their instruction and students' learning, they also believe that students should be accountable for their own learning. The agreement with both improvement and student accountability was also seen in other studies (Brown, 2004; Brown et al., 2009; Brown et al., 2011a; Brown et al., 2011b; Brown & Michalis, 2011; Remesal, 2009; Remesal, 2011; Segers & Tillema, 2011).

Correlation analysis further shows the association between the belief in using assessment for improvement and the belief in making student accountable. A significant correlation was yielded between improvement and student accountability. Significant correlation was also yielded between improvement and school accountability. This finding corroborated Segers and Tillema's (2011) study which revealed significant correlation between improvement and accountability conceptions. These findings indicate that teachers in Belize associate improvement to accountability believing that in order for improvement to occur in teaching and learning, students and schools should be held accountable. Classroom assessment in Belize, as is the case in the Dutch's setting, serve both formative and summative purpose. Teachers use assessment to obtain grades which are then averaged out and used to hold back or promote students to another level. Information received from classroom assessment are also used to plan future instruction or to diagnose students' learning. This is why findings from this study show over eighty five percent of the teachers agreeing with all three items within the student accountability construct; namely placing students into categories, assigning grade to students work and determining if students meet qualification standards. Teachers, over eight five percent, also agreed with nearly all of the items within the improvement conception such as integrating assessment with teaching practices and improving students' learning.

In fact only two items that made up the improvement conception received over twenty percent of disagreement responses. Thirty nine percent of the teachers disagreed with the statement that results are consistent and twenty two percent disagreed that assessment results could be depended on. These disagreement results are consistent with agreement results for similar items within the irrelevancy conception. While majority of the teachers disagreed with items within the irrelevancy conceptions, two items saw high level of agreements. Over ninety percent of the teachers agreed that assessment should be treated cautiously because of measurement error and teachers should take into account the error and imprecision in all. These results are indicating that even though teachers agree with improvement and student accountability conceptions and associate accountability with improvement, they are sceptical about results gained from assessments. This scepticism is seen even more in the fact that a significant correlation was not yielded between improvement and irrelevancy conception. Even though a negative correlation was yielded between improvement and irrelevancy conception indicating that as teachers agreed with improvement conception, they disagreed with irrelevancy, the correlation was not that strong to be considered significant.

The negative correlation between improvement and irrelevant is similar to the results from majority of studies that used Brown's TCoA inventory. Majority of the studies found teachers agreeing with improving conceptions and disagreeing with irrelevancy. As explained by Brown and Hattie (2009), it is natural that teachers who agree with assessment being use for the improvement of teaching and learning will disagree with assessments being irrelevant. However, studies done in Australia, China, Cyprus and Spain showed otherwise. These studies not only saw teachers highly agreeing with improvement, they also highly agreed with irrelevancy (Brown et al., 2011a; Brown et al., 2011b; Brown & Michaelides, 2011; Remesal, 2009). Several reasons are put forth for this such as teachers being aware of inaccuracy in assessment, assessment having no legitimate role in education, error in measurements and the misuse of assessment policies and practices. Even though a negative correlation was yielded in this study between improvement and irrelevancy, that correlation was weak indicating some agreement with irrelevancy. The weakness in correlation might have been due to the items explained earlier. All four items are related to measurement errors suggesting that teachers may believe that assessments are irrelevant due to the measurement errors they had experienced.

6.3.1 Differences in Assessment Conceptions Based on Teachers' Characteristics

Brown (2004) found no significant differences between mean scores for teaching experience and measurement training. While no significant differences were found between assessment conceptions and years of experience in this study, a significant difference was yielded in mean scores between teachers who did not take a university course in assessment / measurement training and those who did for school accountability conception. Teachers who did not take a university course yielded a higher mean score than those who did. This suggests that teachers who have no formal training in measurement and assessment highly believe that assessments should be used to make schools accountable than those who have formal training. This significant result between school accountability and measurement training may be due to the variance in responses for the school accountability conception. School accountability conception yielded the highest standard deviation score compared to the other three conceptions who all yield standard deviation scores of less than one. Responses varied across the six responses with nearly half of the teachers slightly or moderately agreeing with using assessments to make school accountable, a quarter of the teachers mostly or strongly agreeing and a quarter slightly agreeing or mostly disagreeing. Due to this variance in response, a bigger sampling size may have yielded a different result than the significant result yielded between school accountability and measurement training.

While there were slight differences in mean scores for other characteristics such as gender and subject, no statistical significant differences were yielded. Similar mean scores were seen for improvement, student accountability and school accountability conceptions between male and female teachers. There was a slight difference in mean scores for irrelevancy conception with male teachers yielding a higher mean score than female teachers which could be indicating that male teachers believe more than female teachers that assessment is irrelevant to teaching and learning. However, independent *t*-test revealed no significant difference. Similar results were also seen for subject areas. While there was no significant difference in mean scores, science teachers had the lowest mean scores for all conceptions indicating that science teachers' conceptions of assessment vary. These teachers may believe in aspects relating to all four of the conceptions. English teachers yielded the highest mean scores for both accountability conceptions suggesting that these teachers strongly believe that assessment should be used to make students and school accountable. Social studies teachers produced the highest mean scores for improvement and irrelevancy conceptions. These teachers might believe that assessment should be used for improvement, but they also believe that there are errors or inaccuracy in assessment results.

6.4 ASSOCIATION BETWEEN ASSESSMENT CONCEPTIONS AND PRACTICES

Inconsistencies exist regarding the association between teachers' conceptions about assessment and their actual assessment practices. Studies have shown that beliefs held by teachers regarding assessment are in line with their practices (Brown et al., 2009; Calveric, 2010), while others have shown that teachers' beliefs about assessment are out of line with their practices (Chew and Lee, 2013; James and Pedder, 2006). Results from this study are more in line with studies that found a consistent association between beliefs and practices since significant relationships were yielded between conceptions and related practices.

Accountability conceptions were associated with accountability practices. School accountability conception yielded a statistical significance with major exams and assessment provided by publishers, while student accountability conception was significantly associated with using assessment for promotion and graduation, to rank students and to assign grade. These results are in line with earlier results that showed that teachers highly value summative / traditional assessment practices. This might be indicating that what teachers believe, they will practice. Teachers highly conceive assessment as being a tool to make students accountable for their learning; thus, they frequently use assessment practices that are in accordance with their beliefs.

This accountability belief held by teachers are also highlighted in the other two significant associations. Improvement conception was significantly associated with two improvement practices; using information to plan teaching and using information to help students address problems in learning. However, improvement conception was also significantly associated with assessment provided by publishers. These results are in agreement with earlier results that revealed teachers agreeing with both student accountability and improvement conceptions. It is clear that teachers believe that assessment should be used for improvement purposes and also for accountability purposes. This is even clearer with the fact that irrelevancy conception was significantly related with authentic assessment and using assessment without providing grade. This result suggests that teachers not only see the irrelevancy in using authentic assessment, they believe that using assessment without providing a grade is irrelevant.

It is difficult to dismiss these association as statistical errors when they are actually in line with earlier results. From the start it was revealed that teachers used summative assessment practices more frequently than formative assessment practices. It was also revealed that teachers highly agreed with improvement and student accountability conceptions. In fact those two conceptions yielded similar mean scores. The consistent association between improvement and accountability conceptions and related practices was not surprising, but just confirmed the earlier findings.

Studies that concurrently examined assessment conceptions and practices had suggested several reasons for discrepancies between conceptions and practices. Reasons attributed to testing context, assessment system and confusion about assessment have been suggested by studies (Azis, 2015; James & Pedder, 2006). These reasons are based on the premise that even though teachers indicate that they highly value formative practices, they use summative practices more frequently than formative practices. This was not the case in this study. Teachers indicated that they value both formative and summative practices and they indicated that they use both practices; even though they slightly use summative practices more often than formative.

6.5 SUMMARY

Assessment is seen as a driving force in improving teaching and learning. The use of positive assessment in teaching has been shown to improve student’s learning and to equip students with essential 21st century skills. The findings in this study corroborate the thirty years of research findings on assessment practices showing that teachers use summative / tradition assessment practices more often then they use formative assessment practices. One of the reason for this may be teacher's conceptions about assessment since this study also found that teacher's assessment practices are associated with their assessment conceptions. Thus, to change teacher's assessment practices, changes must first be done with teacher's views about assessment. The following chapter details some implications and recommendations based on the major findings discussed in this chapter.

CHAPTER SEVEN
7.0 CONCLUSION

7.1 INTRODUCTION

Classroom assessment has the potential to improve student learning if teachers use the right assessment strategies, provide feedback to students and involve students in the assessment process (Stiggins, 2002). The assessment practices teachers choose to use in their classes have been claimed to be strongly influenced by teacher's beliefs about assessment. As focus turns on assessments; particularly classroom assessment, it is critical not only to know the types of assessment practices frequently used by teachers, but it is also paramount that teacher's conception of assessment is clearly understood. Knowing these will play an important role in implementing national assessment policies. The preceding chapters provided a detailed analysis and discussion of the study's findings, this chapter provides a summary of the key findings, implications of the study and recommendations.

7.2 KEY FINDINGS

This study aimed to examine teachers' assessment conceptions and practices. The target population was secondary teachers who taught English language, mathematics, science and social studies in six government-funded high schools in Belize. Four research questions were used to fulfil the aims of the research. Those were

1. What assessment practices are used by teachers in government funded high schools in Belize?
2. What are secondary school teachers' conceptions of assessment?
3. What are the differences between assessment conceptions and practices and teacher characteristics such as gender, training, teaching experience and subject taught?
4. How do teachers' assessment conceptions relate to their assessment practices?

Even though differences in conceptions and practices were expected, since the Education and Training Act 2010 of Belize dictates that each school should develop their own internal policies and practices for assessing and evaluating students' achievement, the study found that majority of the teachers share similar practices and conceptions pertaining to assessment.

Teachers were found to use a variety of assessment practices and used information from assessment in a variety of ways. Major exams, self-made tests with select-responses, questioning, observation and group projects were discovered to be frequently used assessment practices, while practices such as authentic assessment, assessment from publisher and performance assessment saw the lowest frequency of usages. In regards to the usage of assessment information, teachers' responses indicated that they use a consistent amount of formative and summative practices, but use summative practices more often. For instance, teachers indicated that they frequently used information from assessment to plan teaching, to help student address problems in their learning and to make decisions regarding individual students. However, they also indicated that they frequently used information from assessment for promotion and graduation, to assign grade and to rank students. Furthermore, a high number of teachers indicated that they rarely used assessment without providing a grade.

Teachers' conceptions about assessment were in line with the practices they utilized. They agreed with both improvement and student accountability conceptions. Significant correlations were also yielded between improvement and student accountability and improvement and school accountability indicating that teachers associate improvement with accountability practices.

The mean scores were analysed for differences in responses between assessment conceptions and practices and teacher's gender, experience, subject taught and training. No statistical differences were found between gender and assessment practices and conceptions. Significant difference was yielded between experience and only one assessment practice. Teachers with less than three years of teaching experience were found to use performance assessment less often than teachers with more than four years of teaching experience. Significant differences were also yielded between subject taught and assessment practices and usage of assessment information. Science teachers were found to use tests with select-type items more often than English Language teachers, while mathematics teachers were found to use essays less often than science, social studies and English Language teachers. Math teachers were also found to use information from assessment to plan teaching, to address problems in students' learning and to assign grades more often than teachers in other subject areas. Lastly, statistical significances were also found between training and assessment practices and conceptions. Teachers who took a university course in measurement and assessment were found to use performance assessment more often than those who did not. Teachers who did not take a university course were also found to agree more with school accountability conception than those that took a course.

As it relates to association between conceptions and practices, the study found that what teachers believe about assessment, they practice. Teachers' beliefs about accountability and improvement conceptions were associated with practices such as major exams, assessment provided by publisher, using assessment information for promotion and graduation, to rank students, to plan teaching and to help students address problems in learning.

7.3 IMPLICATION OF THE STUDY

This study aimed to examine teachers' assessment conceptions and practices not only to determine assessment practices commonly used and conceptions that teachers have about assessment, but to examine the association between teachers' conception about assessment and their commonly used assessment practices. Interesting findings have emerged from this study that could potentially affect assessment in Belize and beyond.

7.3.1. Implication for Stakeholders

Majority of the teachers indicated their agreement with using assessment for accountability purposes which suggests that they highly believe that students should be held accountable for assessment results. The significant association between such beliefs and related assessment practices such as using assessment information to rank students and to assign a grade has serious consequences for the different stakeholders in education.

Students are mostly affected from such belief and culture of assessment. When high priority is given to the collection and assigning of grades, it creates students that are more concerned with gaining the right grades rather than gaining the right knowledge. These students may begin to define themselves by the grades which may have serious repercussions especially for low achieving and disabled students. Summative practices such as ranking students have been proven to affect a child's self-efficacy beliefs. Alkharusi (2010) found that students' self-efficacy was positively related to perceived learning assessment environment and negatively related to perceived harsh assessment environment. Students who are frequently assessed using summative / traditional practices may develop a low belief in their ability to succeed in school.

It is encouraging that the findings revealed that the teachers do use some formative practices such as using assessment to plan teaching and to identify students' weakness. Teachers also indicated that they use practices such as observation, questioning and group projects at a high frequency. Even though the effective usages of those practices is uncertain, the fact that they incorporate such practices indicate that teachers are willing to experiment with acclaimed assessment practices once they are introduced to them. The promotion of such acclaimed assessment practices is needed since key alternative assessment practices such as self-assessment, peer assessment, performance assessment and authentic assessments received low rating from the teachers. In fact, the usage of authentic assessment and using assessment without providing grade were statistically associated with irrelevancy conceptions suggesting that teachers found both practices to be irrelevant to the teaching and learning environment.

Such findings reveal the need for effective teacher training program and staff development workshop. Teachers training program may not only affect teachers' assessment practices, but also their assessment belief since independent *t* – test revealed statistical differences between both assessment practice and conception and measurement training. Teachers who received formal training in assessment used practices such as performance assessment and observation more often than those who did not receive formal training. Teachers' belief system may also be impacted by formal training in assessment since teachers who did not take a course in assessment had a greater agreement with using assessment to keep students accountable than those who did take a course.

Teachers' belief about assessment has been impacted not only by their own experiences as a student, but also by the societal and cultural norms. This study found that what teachers believe about assessment impact their assessment practices. Thus, to change teachers' assessment practices attention should firstly be placed on changing their conceptions about assessment. The development of sound and effective teacher training is one way to do that.

Policymakers' attention then should not only be concerned with providing teachers with training and staff development on effective assessment practices, they should first place attention on changing teachers' assessment conception. As Nespor (1987) ascertained, examining teacher's belief is a way to understand how teachers' see their work. Brown-Lopez (2010) noted that primary school teachers' instructional practices were greatly impacted by their teaching and learning experiences. This corroborates Shepard (2000) who illustrated how the summative / traditional beliefs that teachers have towards assessment is popular because it has infiltrated the cultural norms of societies. Teachers, parents and even students view assessments as tests to be taken and grades to be received. Teachers may even find themselves in problems with parents and the society if they choose to assess students mainly through qualitative means rather than quantitative means. Thus, to change the ways teachers assess students, changes need to be made not only to teachers' conception about assessments, but also to society's conception about assessments. Policymakers should then not only ensure that teachers are provided with the relevant assessment training and staff development, but also ensure that policies regarding assessments are known and understood by the wider society.

7.3.2. Implication for Further Studies

Based on the findings and information gathered from the literature, there are several other in-depth research that could be undertaken. A bigger sample that incorporates both public and denominational high schools could be used in a future study. Other research methods such as observation and interviews could be used to support quantitative data gathered. These methods could also be used to examine the effectiveness of teacher's self-reported practices and to gain a better understanding of their view about assessment. This research also provides an avenue for comparative studies between public and denominational / private high schools. Such a study will be able to validate the findings in this study.

The findings revealed a struggle between the usage of formative assessment and summative assessment practices. Teachers simultaneously agreed with using assessment to make students accountable and using it to improve teaching and learning. A future study could examine the usage of formative assessment and the reasons teachers prefer using summative practices over formative practices.

7.4 RECOMMENDATIONS

This study has identified areas of concerns with the practices and beliefs of classroom assessment at the secondary level in Belize. On the basis of the findings in this study, the following recommendations are made:

1. This study is in agreement with Belize Education Sector Strategy 2011 – 2016 which calls for the implementation of a national secondary school curriculum. Before utilizing positive assessment practices, teachers need to be aware of standards being aimed for in each subject area. The fact that teachers indicated that they frequently use summative practice to assess student's learning even though there is no clear national standards is troubling. This may be indicating that teachers use assessment mostly to gain grades to meet the accountability measures within their schools rather than to improve or gauge student' learning. The implementation of a national secondary school curriculum will be the road map for teachers to use to gear

students' achievement through the use of assessments. Teachers can then use assessments to navigate student's learning from one goal to another goal.

2. The Ministry of Education (MOE) and school administration should provide training and certification in formative assessment. The MOE could utilize teacher's professional development workshop in the summers to train teachers in using the different formative assessment practices effectively in their classroom. School administration could use in-service workshops to broaden teacher's understanding of formative assessments, but also to prevail upon teachers the importance and necessity of formative assessments in teaching and learning. These workshops and trainings should not solely be about providing teachers with a list of formative practices, but those programs should help teachers become reflective about their beliefs and practices.

3. Schools can create buddy programs by pairing up novice teachers with experienced teachers. Many times novice teachers are left alone to fend for themselves in the classroom. As was revealed in this study, novice teachers agreed more with using assessment to make students accountable than experienced teachers. Pairing off with experienced teachers will allow novice teachers to not only view assessment as an improvement tool, but it might also make them use formative practices a lot more often than they would have used at the beginning of their career.

4. Teacher's education programs should include at least two courses on classroom assessment. The purpose of these courses should be more than passing on knowledge about the different types of assessment or different measurement strategies. It is vital that courses in classroom assessment take into consideration teachers' existing conceptions about assessment. Providing teachers with knowledge about formative assessment practices will do no good if their assessment beliefs are not confronted. Changes in teacher's assessment practices will only occur once their beliefs are in line with their practices. Thus, teacher's education programs should aim at not only providing information about classroom assessment, but also confront

teachers existing beliefs about assessment. Furthermore, teacher's education program should merge the connection between assessment and instruction. Many teachers use summative practices over formative practices because they believe that assessments cannot be used for instruction. Teacher's education program can alter this belief by promoting the linkage between assessment and instruction.

7.5 SUMMARY

This study looked at secondary teachers' assessment practices and beliefs in Belize. It found that teachers use a variety of assessment practices but tend to favour summative / traditional practices. Teachers' assessment practices were in accordance with their assessment conception since they highly agreed with both student accountability and improvement accountability conceptions. This result suggests that in order to change teachers' assessment practices, their conceptions about assessment needs to be changed.

REFERENCES

Abelson, R. P. (1979). Differences between belief and knowledge systems. *Cognitive Science, 3*(4), 355-366.

Adams, T. L., & Hsu, J. W. Y. (1998). Classroom assessment: Teachers’ conceptions and practices in mathematics. *School Science and Mathematics, 98*(4), 174-180.

Alkharusi, H. (2010). A multilevel linear model of teachers’ assessment practices and students’ perceptions of the classroom assessment environment. *Procedia-Social and Behavioral Sciences, 5*, 5-11.

Alkharusi, H. (2011). Teachers' classroom assessment skills: Influence of gender, subject area, grade level, teaching experience and in-service assessment training. *Journal of Turkish Science Education, 8*(2).

Alkharusi, H., Aldhafri, S., Alnabhani, H., & Alkalbani, M. (2014). Classroom assessment: Teacher practices, student perceptions, and academic self-efficacy beliefs. *Social Behavior and Personality: An International Journal, 42*(5), 835-855.

Allahyar, N., & Nazari, A. (2012). Potentiality of Vygotsky's sociocultural theory in exploring the role of teacher perceptions, expectations and interaction strategies. *WoPaLP, 6*, 79-92.

Andrade, H., & Cizek, G. J. (2010). *Handbook of formative assessment*: New York, NY: Routledge.

Armstrong, C. J. (2011). *Belizean primary school teachers' understanding of assessment; assessment practices, and use of student assessment data* (Doctoral Dissertation). Retrieved from http://search.proquest.com/docview/924514273?accountid=28103 ProQuest Central database.

Ashton, P. (2015). Historical overview and theoretical perspectives of research on teachers’ beliefs. In H. Fives & M. Gill (eds.), *International Handbook of research on teachers’ beliefs* (pp. 31-47). New York, NY: Routledge.

Azis, A. (2012). Teachers' conceptions and use of assessment in student learning. *Indonesian Journal of Applied Linguistics, 2*(1), 40-52.

Azis, A. (2015). Conception and practices of assessment: A case of teachers representing improvement conception *TEFLIN Journal, 26*(2), 129-154.

Babbie, E. (2015). *The practice of social research*. Boston, MA: Cenagage.

Barnes, N., Fives, H., & Dacey, C. M. (2015). Teachers' beliefs about assessment. In H. Fives & M. Gill (eds.), *International handbook of research on teachers' beliefs* (pp. 284-300). New York, NY: Routledge.

Barnett, C., Catzim-Sanchez, A., & Humes, D. (2010). *Report of the national consultations on horizon 2030.* Belize City, BZ: Barnett & Co. Ltd.

Beckett, D., Volante, L., & Drake, S. (2010). Formative assessment: Bridging the research-practice divide. *Education Canada, 50*(3), n3.

Bennett, J. A. (2008). *Education in Belize: A historical perspective*. Belize, BZ: Angelus Press.

Berry, R., & Adamson, B. (2011). Assessment reform past, present and future. In R. Berry & B. Adamson (Eds.), *Assessment reform in education: Policy and practice* (pp. 3-14). New York, NY: Springer.

Black, Paul; Harrison, Christine; Lee, Clare; Marshal, Bethan and Wiliam, Dylan (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86(1) pp. 8–21.

Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7-74. http://dx.do[i]i.org/10.1080/0969595980050102

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education), 21*(1), 5-31.

Bol, L., Stephenson, P. L., O'connell, A. A., & Nunnery, J. A. (1998). Influence of experience, grade level, and subject area on teachers' assessment practices. *The Journal of Educational Research, 91*(6), 323-330.

Broadfoot, P., Daugherty, R., Gardner, J., Harlen, W., James, M., & Stobart, G. (2002). *Assessment for learning: 10 principles*. Cambridge, UK: University of Cambridge School of Education.

Broadfoot, P., & Black, P. (2004). Redefining assessment? The first ten years of assessment in education. *Assessment in Education, 11*(1), 7-26.

Broadfoot, P. (2007). *An introduction to assessment*. London: Continuum.

Brookhart, S. (2004). Classroom assessment: Tensions and intersections in theory and practice. *The Teachers College Record, 106*(3), 429-458.

Brookhart, S. (2007). Expanding views about formative classroom assessment: A review of the literature. In J. H. McMillan (Ed.), *Formative classroom assessment: Theory into practice* (pp. 43-62). New York, NY: Teachers College Press.

Brown, G. T. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice, 11*(3), 301-318.

Brown, G. T. (2006). Teachers' conceptions of assessment: Validation of an abridged version 1, 2. *Psychological Reports, 99*(1), 166-170.

Brown, G. T., & Gao, L. (2015). Chinese teachers' conceptions of assessment for and of learning: Six competing and complementary purposes. *Cogent Education, 2*(1), 993836.

Brown, G. T., Hui, S. K., Flora, W., & Kennedy, K. J. (2011a). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research, 50*(5), 307-320.

Brown, G. T., Kennedy, K. J., Fok, P. K., Chan, J. K. S., & Yu, W. M. (2009). Assessment for student improvement: Understanding Hong Kong teachers' conceptions and practices of assessment. *Assessment in Education: Principles, Policy & Practice, 16*(3), 347-363.

Brown, G. T. L. (2003). *Teachers' instructional conceptions: Assessment's relationship to learning, teaching, curriculum, and teacher efficacy*. Paper presented to the joint conference of the Australian and New Zealand Associations for Research in Education (AARE/NZARE), Auckland, NZ, November 28–December 3, 2003.

Brown, G. T. L. (2011). Teachers' conceptions of assessment: Comparing primary and secondary teachers in New Zealand. *Assessment Matters, 3*, 45-70.

Brown, G. T. L., & Harris, L. R. (2013). Student self-assessment. In J. H. McMillan (Ed.). *The SAGE handbook of research on classroom assessment* (pp. 367-393). Thousand Oaks, CA: Sage

Brown, G. T. L., & Hattie, J. A. C. (2009). *Understanding teachers' thinking about assessment: Insights for developing better educational assessment*. Paper presented at the Annual Meeting of the National Council for Measurement in Education (NCME), San Diego, CA, April 2009 Retrieved from http://repository.lib.ied.edu.hk/pubdata/ir/link/pub/2009%20Brown%20and%20Hattie%20NCME%20v4.pdf

Brown, G. T. L., Lake, R., & Matters, G. (2011b). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and Teacher Education, 27*(1), 210-220. doi: http://dx.doi.org/10.1016/j.tate.2010.08.003

Brown, G. T. L., & Michaelides, M. P. (2011). Ecological rationality in teachers' conceptions of assessment across samples from Cyprus and New Zealand. *European Journal of Psychology of Education, 26*(3), 319-337.

Brown-Lopez, P. (2010). *Analysis of the effects of a constructivist-based mathematics problem solving instructional program on the achievement of grade five students in Belize, Central America* (Doctoral Dissertation). Retrieved from http://etheses.dur.ac.uk/482/1/P._Lopez_Thesis_Pdf.pdf?DDD29+

Buehl, M. M., & Beck, J. S. (2015). The relationship between teachers' beliefs and teachers' practices. In H. Fives & M. Gill (eds.), *International handbook of research on teachers' beliefs* (pp. 66-84). New York, NY: Routledge.

Calveric, S. (2010). *Elementary teachers' assessment beliefs and practices* (Doctoral Dissertation). Retrieved from http://scholarscompass.vcu.edu/cgi/viewcontent.cgi?article=3331&con

Carneiro, F. G. (2016). *Belize, right choices bright future: systematic country diagnostic*. Washington, D.C: World Bank Group. Retrieved from http://documents.worldbank.org/curated/en/870551467995073017/Belize-right-choices-bright-future-systematic-country-diagnostic

Chase, C. (1999). *Classroom assessment for educators*. New York, NY: Addison-Wesley

Chen, C.-H. (2008). Why do teachers not practice what they believe regarding technology integration?. *The Journal of Educational Research, 102*(1), 65-75.

Chew, A. Y. L., & Lee, I. C. H. (2013). *Teachers' beliefs and practices of classroom assessment in Republic Polytechnic, Singapore*. Retrieved from http://www.iaea.info/documents/paper_5b92a581.pdf

Cizek, G. J. (2010). An introduction to formative assessment: History, characteristics, and challenges. In G. J. Cizek and H. L. Andrade (Eds.), *Handbook of formative assessment* (pp. 3-17). New York, NY: Routledge.

Cizek, G. J., Fitzgerald, S. M., & Rachor, R. A. (1995). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational assessment, 3*(2), 159-179.

Clark, I. (2010). Formative assessment: 'There is nothing so practical as a good theory'. *Australian Journal of Education*, *54*(3), 341-352. http://dx.doi.org/10.1177/000494411005400308

Clark, I. (2012). Formative assessment: assessment is for self-regulated learning. *Educational Psychology Review, 24*(2), 205-249. Retrieved from http://www.jstor.org/stable/43546787

Collins, H. (2010). *Creative research: the theory and practice of research for the creative industries*. London, UK: Ava Publishing.

Creswell, J. (2012). *Educational research: Planning, conducting and evaluating quantitative and qualitative research*. 4th ed. Boston: Peason.

Cross, R. (2010). Language teaching as sociocultural activity: Rethinking language teacher practice. *The Modern Language Journal, 94*(3), 434-452.

Crossouard, B., & Pryor, J. (2012). How theory matters: Formative assessment theory and practices and their different relations to education. *Studies in Philosophy and Education, 31*(3), 251-263.

Dayal, H. & Lingam, G. (2015). Fijian teachers' conceptions of assessment. *AJTE*, *40*(40). http://dx.doi.org/10.14221/ajte.2015v40n8.3

DeLuca, C., & Lam, C. Y. (2014). Preparing teachers for assessment within diverse classrooms: an analysis of teacher candidates' conceptualizations. *Teacher Education Quarterly, 41*(3), 3.

Di Gropello, E. (2003). *Monitoring educational performance in the Caribbean (Vol. 6)*. Washington, DC: World Bank Publications.

Duncan, C. R., & Noonan, B. (2007). Factors affecting teachers' grading and assessment practices. *Alberta Journal of Educational Research, 53*(1), 1.

Erkmen, B. (2012). Ways to uncover teachers' beliefs. *Procedia - Social and Behavioral Sciences, 47*, 141-146. doi: http://dx.doi.org/10.1016/j.sbspro.2012.06.628

Ferrer, J. G. (2006). *Educational assessment systems in Latin America: Current practice and future challenges*. Washington, DC: PREAL.

Fives, H. & Gill, M. (Eds.). (2015). *International handbook of research on teachers' beliefs*. New York: Routledge

Freire, P. (1970). *Pedagogy of the oppressed*. New York, NY: Herder and Herder.

Frey, B. B., & Schmitt, V. L. (2007). Coming to terms with classroom assessment. *Journal of Advanced Academics*, 18(3), 402-423, doi:10.4219/jaa-2007-495.

Frey, B. B., & Schmitt, V. L. (2010). Teachers' classroom assessment practices. *Middle Grades Research Journal, 5*(3), 107-117.

Gayle, H. and Mortis, N. (2010). *Male social participation and violence in urban Belize: An examination of their experience with goals, guns, gangs, gender, god, and governance*. Belize City, BZ. Retrieved from http://dbzchild.org/uploads/docs/complete_pgmale_social_participation_and_violence_in_urban_belize_grand.pdf

Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment*. London: The Falmer Press.

Gipps, C. (1999). Socio-Cultural aspects of assessment. *Review of Research in Education*, *24*, 355. http://dx.doi.org/10.2307/1167274

Glasser, T. (1990). *The quality school*. New York, NY: Harper & Row.

Government of Belize, *Education Act*, Chapter 36, (December 31, 2000), [cited March 17, 2016]; available from http://www.belizelaw.org/web/lawadmin/index2.html

Gredler, M. (1999). *Classroom assessment and learning*. New York, NY: Longman.

Gullickson, A. R. (1985). Student evaluation techniques and their relationship to grade and curriculum. *The Journal of Educational Research, 79*(2), 96-100.

Halcrow Group Limited. (2010). *Country poverty assessment report*. Belize.

Government of Belize and the Caribbean Development Bank. Retrieved from http://www.caribank.org/uploads/2012/12/Belize-2009-Report-Vol1.pdf

Hargreaves, A., Earl, L., & Schmidt, M. (2002). Perspectives on alternative assessment reform. *American Educational Research Journal*, *39*(1), 69-95. http://dx.doi.org/10.3102/00028312039001069

Harlen, W. (1994). *Enhancing quality in assessment*. London: P. Chapman Pub.

Harlen, W. (2005). Teachers' summative practices and assessment for learning–tensions and synergies. *Curriculum Journal, 16*(2), 207-223.

Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112. http://dx.doi.org/10.3102/003465430298487

Heritage, M. (2007). Formative assessment: What do teachers need to know and do?. *Phi Delta Kappan*, *89*(2), 140-145. http://dx.doi.org/10.1177/003172170708900210

Hills, J. R. (1991). Apathy concerning grading and testing. *Phi Delta Kappan, 72*(7), 540-545.

Hughes, G. B. (2010). Formative assessment practices for students at risk. In H. L. Andrade and G. J. Cizek (Eds.). *Handbook of formative assessment (*pp. 212-232). Madison Avenue, NY: Routledge

James, M., & Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *The Curriculum Journal, 17*(2), 109-138.

John-Steiner, V., & Mahn, H. (1996). Sociocultural approaches to learning and development: A Vygotskian framework. *Educational psychologist, 31*(3-4), 191-206.

Kagan, D. M. (1992). Implication of research on teacher belief. *Educational psychologist, 27*(1), 65-90.

Kingston, N., & Nash, B. (2011). Formative Assessment: A Meta-Analysis and a Call for Research. *Educational Measurement: Issues and Practice, 30*(4), 28-37. doi: 10.1111/j.1745-3992.2011.00220.x

Krause, K., Bochner, S., Duchesne, S., & McMaugh, A. (2010). *Educational psychology for learning and teaching* (3rd ed.). Melbourne: Cangage Learning Australia.

Li, W. S. & HUI, K. F. S. (2007). Conceptions of assessment of Mainland China college lecturers: A technical paper analyzing the Chinese version of COA-III. *The Asia Pacific-Education Researcher*, 16(2), 185-198. Retrieved from http://www.dlsu.edu.ph/research/journals/taper/pdf/200712/Li%20&%20Hui.pdf

Lodico, M., Spaulding, D., & Voegtle, K. (2006). *Methods in educational research*. San Francisco, CA: Jossey-Bass.

Looney, J. W. (2011). Integrating formative and summative assessment: Progress toward a seamless system. *OECD Education Working Papers*, No. 58, OECD Publishing. http://dx.doi.org/10.1787/5kghx3kbl734-en

Marso, R. N., & Pigge, F. L. (1991). An analysis of teacher-made tests: Item types, cognitive demands, and item construction errors. *Contemporary Educational Psychology, 16*(3), 279-286.

Martell, K. (2007). Assessing student learning: Are business schools making the grade?. *Journal of Education for Business*, *82*(4), 189-195. http://dx.doi.org/10.3200/joeb.82.4.189-195

McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice, 20*(1), 20-32.

McMillan, J. (2013). Why we need research on classroom assessment. In J. H. McMillan (Ed.). *The SAGE handbook of research on classroom assessment* (pp. 3-16). Thousand Oaks, CA: Sage

McMillan, J. H., Cohen, J., Abrams, L., Cauley, K., Pannozzo, G., & Hearn, J. (2010). Understanding secondary teachers' formative assessment practices and their relationship to student motivation. *Online Submission*. Retrieved from http://eric.ed.gov/?id=ED507712

McMillan, J. H., Venable, J. C., & Varier, D. (2013). Studies of the effect of formative assessment on student achievement: So much more is needed. *Practical Assessment, Research & Evaluation, 18*(2), 2.

Mertler, C. A. (1998). Classroom Assessment Practices of Ohio Teachers. Paper presented at the meeting of the Mid-Western Educational Research Association, Chicago.

Mertler, C. A. (2003). *Classroom assessment: A practical guide for educators*. Los Angeles, CA: Pyrczak

Ministry of Education. (2012). *Belize Education Sector Strategy 2011 - 2016*. [online] Belize. Retrieved from: http://www.moe.gov.bz/index.php/downloads/category/19-education-sector-strategy

Ministry of Education. (2015a). *Abstract of education statistics*. [online] Belize. Retrieved from: http://www.moe.gov.bz

Ministry of Education. (2015b). Primary school examination preliminary results 2015 [Press release]

Moss, C. (2013). Research on classroom summative assessment. In J. H. McMillan (Ed.). *The SAGE handbook of research on classroom assessment* (pp. 235-256). Thousand Oaks, CA: Sage

Näslund-Hadley, E., Alonzo, H., & Martin, D. (2013). *Challenges and opportunities in the Belize education sector*. Inter-American Development Bank. Retrived from https://publications.iadb.org/handle/11319/5926?locale-attribute=en

Ndalichako, J. L. (2015). Secondary school teachers' perceptions of assessment. *International Journal of Information and Education Technology, 5*(5), 326.

Ndwandwe, S. B., & Dlamini, B. M. (2014). A descriptive study of high school agriculture teacher's competencies in Swaziland. *NACTA Journal*, 58(2), 88-94. Retrieved from http://search.proquest.com/docview/1537034606?accountid=28103

Nespor, J. (1987). The role of beliefs in the practice of teaching. *Journal of curriculum studies, 19*(4), 317-328.

Neuman, W.L. (2006). Social Research Methods: Qualitative and Quantiative Approaches. Toronto: Pearson.

Nitko, A.J. (2004). *Educational assessment of students (4th ed)*. Ohio: Merrill Prentice Hall

Noonan, B., & Duncan, C. R. (2005). Peer and self-assessment in high schools. *Practical Assessment, Research and Evaluation, 10*(17), 1-8.

Ogunkola, B. J., & Clifford, C. (2013). Instructional assessment practices of science teachers in Barbados: Pattern, techniques and challenges. *Academic Journal of Interdisciplinary Studies, 2*(1), 313.

Ohlsen, M. T. (2007). Classroom assessment practices of secondary school members of NCTM. *American Secondary Education*, 4-14.

Organization for Economic Co-operation and Development. (2005). *Formative Assessment: Improving Learning in Secondary Classrooms*. Paris: Organization for Economic Co-operation and Development (OECD).

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62*(3), 307-332.

Pittaway, L. & Edwards, C. (2012). Assessment: Examining practice in entrepreneurship education. *Education + Training*, *54*(8/9), 778-800. http://dx.doi.org/10.1108/00400911211274882

Popham, W. (2005). *Classroom Assessment*. Boston: Pearson/Allyn and Bacon.

Rashidi, N., & Moghadam, M. (2015). The Discrepancy between teachers' belief and practice, from the sociocultural perspective. *Studies in English Language Teaching, 3*(3), 252.

Remesal, A. (2009). *Spanish students teachers' conceptions of assessment when starting their career*. Paper presented at the Perceptions and conceptions of assessment in the classroom: Different national perspectives, Amsterdam

Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education, 27*(2), 472-482.

Resnick, J. (1989). Inadequate test of knowledge. *BMJ*, *299*(6696), 460-460. http://dx.doi.org/10.1136/bmj.299.6696.460-a

Robinson, L., Maldonado, N., & Whaley, J. (2014). Perceptions about implementation of differentiated instruction. In *The Annual Mid-South Education Research (MSERA) Conference*. Knoxville, Tennessee. Retrieved from http://files.eric.ed.gov/fulltext/ED554312.pdf

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional science, 18*(2), 119-144.

Scott, S., Webber, C. F., Aitken, N., & Lupart, J. (2011). *Developing teachers' knowledge, beliefs, and expertise: Findings from the Alberta Student Assessment Study.* Paper presented at the The Educational Forum.

Segers, M., & Tillema, H. (2011). How do Dutch secondary teachers and students conceive the purpose of assessment?. *Studies in Educational Evaluation, 37*(1), 49-54. doi: http://dx.doi.org/10.1016/j.stueduc.2011.03.008

Serafini, F. (2000). Three paradigms of assessment: Measurement, procedure, and inquiry. *The Reading Teacher, 54*(4), 384-393.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational researcher, 29*(7), 4-14.

Shepard, L. A. (2013). Foreword. In J. H. McMillan (Ed.). *The SAGE handbook of research on classroom assessment* (pp. xix). Thousand Oaks, CA: Sage

Skott, J. (2015). The promises, problems, and prospects of research on teachers' beliefs. In H. Fives & M. Gill (eds.), International *handbook of research on teachers' beliefs* (pp. 13-30). New York, NY: Routledge.

Song, E., & Koh, K. (2010). Assessment for learning: Understanding teachers' beliefs and practices. *IAEA, Assessment for the future generation*, retrieved from http://www.iaea2010.com/index.php?option=com_content&view=article&id=111:assessment-for-learning-understanding-teachers-beliefs-and-practices&catid=37:classroom-assessment&Itemid=89

Statistical Institute of Belize. (2013). *Abstract of statistics*. Belize. Retrieved from: http://www.sib.org.bz/

Stiggins, R. (2014). Improve assessment literacy outside of schools too. *Phi Delta Kappan, 96*(2), 67-72.

Stiggins, R., & Chappuis, J. (2005). Using student-involved classroom assessment to close achievement gaps. *Theory into Practice, 44*(1), 11-18.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan, 83*(10), 758.

Stiggins, R. J., & Bridgeford, N. J. (1985). The ecology of classroom assessment. *Journal of Educational Measurement, 22*(4), 271-286.

Stiggins, R. J., & Chappuis, J. (2012). *An introduction to student-involved assessment for learning*. Boston, MA: Pearson.

Stiggins, R. J., & Conklin, N. F. (1992). *In teachers' hands: Investigating the practices of classroom assessment*. SUNY Press.

Suah, S. L., & Ong, S. L. (2012). Investigating assessment practices of in-service teachers. *International Online Journal of Educational Sciences, 4*(1), 91-106.

Taras, M. (2005). Assessment–summative and formative–some theoretical reflections. *British Journal of Educational Studies, 53*(4), 466-478.

Thomasenia Lott, A., & Jin-Wen Yang, H. (1998). Classroom assessment: Teachers's conceptions and practices in mathematics. *School Science and Mathematics, 98*(4), 174-180.

Thompson, A. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127-146). New York: Macmillan

Thompson, C. (2000). Education reform in post war post-independence: Education in Belize towards the year 2000. Belize, BZ: Spear.

Tierney, R. D. (2006). Changing practices: Influences on classroom assessment. *Assessment in Education, 13*(3), 239-264.

Topping, K. (2013). Peer as a source of formative and summative assessment. In J. H. McMillan (Ed.). *The SAGE handbook of research on classroom assessment* (pp. 395-41). Thousand Oaks, CA: Sage

Torrance, H., & Pryor, J. (1998). *Investigating formative assessment: Teaching, learning and assessment in the classroom*. London, UK: McGraw-Hill Education.

Van den Berg, R. (2002). Teachers' meanings regarding educational practice. *Review of Educational Research, 72*(4), 577-625.

Volante, L., & Beckett, D. (2011). Formative assessment and the contemporary classroom: Synergies and tensions between research and practice. *Canadian Journal of Education, 34*(2), 239-255.

Volante, L., Drake, S., & Beckett, D. (2010). Formative assessment: Bridging the research-practice divide. *Education Canada*, 50(3). Retrieved September 25, 2016 from http://www.cea-ace.ca/education-canada/article/formative-assessment-bridging-research-practice-divide

Ward, A. W., & Murray-Ward, M. (1999). Assessment in the classroom. London: Wadsworth

Wiley, C. R. H. (2011). *Profiles of teacher grading practices: Integrating teacher beliefs, course criteria, and student characteristics* (Doctoral Dissertation). Retrieved from http://search.proquest.com/docview/887719048?accountid=28103 ProQuest Central database.

Wiliam, D. (2011). What is assessment for learning?. *Studies in Educational Evaluation, 37*(1), 3-14.

Wiliam, D., Lee, C., Harrison, C., & Black, P. (2010). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education*.

Wiliam, D. (2013). Feedback and instructional correctives. In J. H. McMillan (Ed.). *The SAGE handbook of research on classroom assessment* (pp. 197-214). Thousand Oaks, CA: Sage

Wolff, L., & Castro, C. d. M. (2000). Secondary education in Latin America and the Caribbean: The challenge of growth and reform. Inter-American Development Bank.

World Bank Group. (2015). *World Development Report 2015: Mind, Society, and Behavior*. Washington, DC: World Bank.

World Bank Group. (2016). *Global economic prospects, January 2016: Spillovers amid weak growth*. Washington, DC: World Bank. doi:10.1596/978-1-4648-0675-9.

Young, V. M., & Kim, D. H. (2010). Using assessments for instructional improvement: A literature review. *Archivos Analíticos de Políticas Educativas= Education Policy Analysis Archives, 18*(19), 1.

Zhang, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education, 16*(4), 323-342.

APPENDICES

APPENDIX A: TEACHER'S QUESTIONNAIRE

Secondary School Teachers' Assessment Beliefs and Practices

Thank you for taking the time to complete this survey. It is based on your classroom assessment practices and your beliefs about assessment. In order to ensure anonymity, please do not put your name or the name of your school on the survey. By completing this survey, it is understood that you do so voluntarily and that you consent to the use of your response in the study.

The survey consists of three parts:

__Part 1__: Asks about your beliefs and understanding about assessment, whatever that term means to you, using a 6-points scale ranging from strongly disagree to strongly agree. Please answer the questions using your own understanding of assessment.

__Part 2__: Asks about the frequency in which you use various types of assessment in your classroom, on a 6-point scale ranging from not at all to very frequently. Please base your response on one of the following courses: Math, English, Science, or Social Studies.

__Part 3__: Demographics

Part 1: Using your own understanding of assessment, choose __one response__ that comes __closest__ to describing your opinion.

Strongly Disagree (1) ☐ Mostly Disagree (2) ☐ Slightly Agree (3) ☐ Moderately Agree (4) ☐ Mostly Agree (5) ☐ Strongly Agree (6) ☐

Conceptions of Assessment	Strongly Disagree	Mostly Disagree	Slightly Agree	Moderately Agree	Mostly Agree	Strongly Agree
1. Assessment provides information on how well schools are doing	☐	☐	☐	☐	☐	☐
2. Assessment places students into categories	☐	☐	☐	☐	☐	☐
3. Assessment is a way to determine how much students have learned from teaching	☐	☐	☐	☐	☐	☐
4. Assessment provides feedback to students about their performance	☐	☐	☐	☐	☐	☐
5. Assessment is integrated with teaching practice	☐	☐	☐	☐	☐	☐
6. Assessment results are trustworthy	☐	☐	☐	☐	☐	☐
7. Assessment forces teachers to teach in a way against their beliefs	☐	☐	☐	☐	☐	☐
8. Teachers conduct assessments but make little use of the results	☐	☐	☐	☐	☐	☐
9. Assessment results should be treated cautiously because of measurement error	☐	☐	☐	☐	☐	☐
10. Assessment is an accurate indicator of a school's quality	☐	☐	☐	☐	☐	☐
11. Assessment is assigning a grade or level to students work	☐	☐	☐	☐	☐	☐
12. Assessment establishes what students have learned	☐	☐	☐	☐	☐	☐
13. Assessment feeds back to student their learning needs	☐	☐	☐	☐	☐	☐
14. Assessment information modifies ongoing teaching of students	☐	☐	☐	☐	☐	☐

Conceptions of Assessment	Strongly Disagree	Mostly Disagree	Slightly Agree	Moderately Agree	Mostly Agree	Strongly Agree
15. Assessment results are consistent	☐	☐	☐	☐	☐	☐
16. Assessment is unfair to student	☐	☐	☐	☐	☐	☐
17. Assessment results are filed and ignored	☐	☐	☐	☐	☐	☐
18. Teachers should take into account the error and imprecision in all assessment	☐	☐	☐	☐	☐	☐
19. Assessment is a good way to evaluate a school	☐	☐	☐	☐	☐	☐
20. Assessment determines if students meet qualifications standards	☐	☐	☐	☐	☐	☐
21. Assessment measures students' higher order thinking skills	☐	☐	☐	☐	☐	☐
22. Assessment helps students improve their learning	☐	☐	☐	☐	☐	☐
23. Assessment allows different students to get different instruction	☐	☐	☐	☐	☐	☐
24. Assessment results can be depended on	☐	☐	☐	☐	☐	☐
25. Assessment interferes with teaching	☐	☐	☐	☐	☐	☐
26. Assessment has little impact on teaching	☐	☐	☐	☐	☐	☐
27. Assessment is an imprecise process	☐	☐	☐	☐	☐	☐

Part 2: Please rate how often you have used the following classroom assessment practices based on the subject you identified. Please rate based on the following rating scale. Tick √ ONE only.

Not at All (1)	Very Rarely (2)	Rarely (3)	Occasionally (4)	Frequently (5)	Very Frequently (6)
☐	☐	☐	☐	☐	☐

A. Your assessment practices

Assessment Practices	Not at All	Very Rarely	Rarely	Occasionally	Frequently	Very Frequently
28. Students' self-assessments	☐	☐	☐	☐	☐	☐
29. Individual projects	☐	☐	☐	☐	☐	☐
30. Major exams (e.g. end of term exams, final exams)	☐	☐	☐	☐	☐	☐
31. Observations	☐	☐	☐	☐	☐	☐
32. Group projects	☐	☐	☐	☐	☐	☐
33. Self-made test with select-type items (e.g. multiple choice, true / false, matching)	☐	☐	☐	☐	☐	☐
34. Oral presentations	☐	☐	☐	☐	☐	☐
35. Assessment provided by publishers (e.g. in textbooks)	☐	☐	☐	☐	☐	☐
36. Portfolio (e.g. collection of assignments, work samples)	☐	☐	☐	☐	☐	☐
37. Peer assessment	☐	☐	☐	☐	☐	☐
38. Questioning	☐	☐	☐	☐	☐	☐

2

39. Authentic assessments (e.g. 'real world' performance task)	☐	☐	☐	☐	☐	☐
40. Self-made test with supply-type items (e.g. short answers)	☐	☐	☐	☐	☐	☐
41. Essay type-questions	☐	☐	☐	☐	☐	☐
42. Performance assessment (structure teacher observation)	☐	☐	☐	☐	☐	☐

B. How you use assessment information

Usage of Assessment Information	Not at All	Very Rarely	Rarely	Occasionally	Frequently	Very Frequently
43. To make decision about individual student	☐	☐	☐	☐	☐	☐
44. For students' promotion and graduation	☐	☐	☐	☐	☐	☐
45. To plan teaching	☐	☐	☐	☐	☐	☐
46. To help student address problems in learning	☐	☐	☐	☐	☐	☐
47. To assign grade	☐	☐	☐	☐	☐	☐
48. To provide descriptive feedback to individual student	☐	☐	☐	☐	☐	☐
49. To rank students	☐	☐	☐	☐	☐	☐
50. Use assessment without providing grade	☐	☐	☐	☐	☐	☐

Part 3: Demographics

51. What is your sex?
☐ Male ☐ Female

52. In responding to this survey, I am doing so for (Select One):
☐ Math ☐ Science ☐ English Language ☐Social Studies

53. How many years have you been teaching?
☐ Less than 3 ☐ Between 11 and 20
☐ Between 4 and 10 ☐ More than 20

54. Have you taken any university courses in classroom assessment (for example: Measurement and Evaluation)?
☐ Yes ☐ No

55. Have you received any in-service / workshop training in assessment, tests and measurement?
☐ Yes ☐ No

56. What is your educational qualifications?
☐ Certificate
☐ Diploma
☐ BS / BA Degree
☐ Master's Degree
☐ Other (Please specify):__________________________

Thanks for taking the time to complete this questionnaire!

3

APPENDIX B: APPROVAL OF HUMAN ETHICS FORM

USP
FACULTY OF ARTS, LAW AND EDUCATION

The University of the South Pacific
Private Mail Bag, Laucala Campus
Suva, Fiji

Ph: (676) 323 2742
Fax: (676) 323 1552
www.usp.ac.fj

FALE 10/16

Gina Middleton
Student ID#: S11119266
School of Education
Faculty of Arts, Law & Education
The University of the South Pacific

6th July 2016

Re: Human Ethics Application

The human ethics application for the following research project has been approved by the FALE Research Committee.

Titled:	Assessment Conception and Practices of Teachers in Public High Schools in the Belize District
Principal Researcher:	Gina Middleton
Program:	Master of Arts in Education
School/Division:	School of Education

The ethical values and principles apply to all University activities, to all its staff and student researchers including those visiting for short periods, and to any research agreements or partnerships that the University establishes. The University's human ethics will be compliant with the laws of individual University member states, particularly in relation to privacy, confidentiality, ownership, intellectual property requirements, research permit requirements and human rights.

All research conducted by persons affiliated with the University of the South Pacific will be carried out only with the prior, free, and informed consent of all persons concerned, whether individuals or communities, based on adequate information. The consent may be withdrawn by a particular individual or community at any time for any reason without disadvantage or prejudice.

Where research involving human subjects/participants is proposed by a researcher at the University of the South Pacific, agreement must be obtained from a fraction of those subjects or their spokespersons in advance of the proposal being submitted to the University Research Ethics Committee, who will require written evidence of such an agreement. Where appropriate, this must be translated into a language that is readily comprehended by potential participants. The rights of participants must be emphasised, their questions all satisfactorily answered.

Please note that it is the researcher's responsibility to submit relevant application forms for Research permit purposes. These must be submitted as annexes with the first progress report form. Also, be reminded that progress reports are expected at end of each six-month period throughout the duration of the project.

Congratulations and all the best for your research project.

Kedrayate

Dr Akanisi Kedrayate
Dean - Faculty of Arts, Law & Education
The University of the South Pacific
Laucala Campus, Suva.

APPENDIX C: LETTER TO GENERAL MANAGER

26 April 2016

Mr. Marvin Kellyman
General Manager of Government Schools
North Front Street
Belize City, Belize

Dear Mr. Kellyman,

REQUEST FOR PERMISSION TO CONDUCT RESEARCH

As part of the requirements of the University of South Pacific's Master of Arts in Education programme, I am conducting research for the purpose of analysing secondary teachers' classroom assessment beliefs and practices.

I am hereby seeking your consent to administer a questionnaire to government high school teachers within the six government high schools in the district of Belize. The questionnaire will allow them to indicate their conceptions of assessment and their current assessment practices. To assist you in reaching a decision, I have attached a copy of the research instrument which I intend using in my research

Should you require any further information, please do not hesitate to contact me or my supervisor. Our contact details are as follows:

Dr. Ledua Waqailiti – Phone #: 679-323-2296 or Email: ledua.waqailiti@usp.ac.fj.
Gina Middleton – Email: s11119266@student.usp.ac.fj

Upon completion of the study, I undertake to provide you with a bound copy of the research study.

Your permission to conduct this study will be greatly appreciated.

Yours sincerely,

G Middleton

Gina Middleton
Master's Candidate
University of the South Pacific

APPENDIX D: LETTER TO PRINCIPALS

2 May 2016

Dear Principal,

REQUEST FOR PERMISSION TO CONDUCT RESEARCH

As part of the requirements of University of the South Pacific's Master of Arts in Education programme, I am conducting research for the purpose of analysing secondary teachers' classroom assessment beliefs and practices.

The research is aimed at secondary teachers within six government schools in the Belize district. I would welcome your school's participation in this 15 minutes survey. The study is limited only to English, Math, Science and Social Studies teachers. Each teacher's participation is entirely voluntary. The promise of strict confidentiality is assured in both the collection and reporting of the data. Any findings obtained in connection with this study will be presented in such a way that no individual school or person will be identifiable. By completing this survey, your teachers will be giving me permission to publish aggregated results in my thesis.

As a fellow secondary school teacher, I am hopeful that the study's findings will assist with providing a list of common assessment practices among high school teachers and clearly defining how teachers' assessment beliefs relate to their classroom assessment practices. Understanding teachers' assessment practices and their beliefs can help with the formulation of professional development that will improve teachers' assessment pedagogies and practices which will enhance instructional planning and student learning.

Should you approve your teachers' participation in this research study, Miss. Trycelin Lord (a teacher) will visit your school between May 2 – 20, 2016 (or a date convenient to you) to administer the survey to selected teachers. This visit will occur after regular classes have ended to provide teachers with the appropriate time to complete the survey.

Please feel free to review the attached survey instrument. Should you have any questions about this study, please contact me at s11119266@student.usp.ac.fj.

Thank you in advance for your time and consideration. This study could not be completed without your help.

Sincerely,

G Middleton

Gina Middleton
Master's Candidate
University of the South Pacific

APPENDIX E: LETTER TO RESEARCH ASSISTANT

May 3, 2016

Dear Miss Trycelin Lord,

Thanks for agreeing to assist me in this crucial part of my research study. The following instructions are provided to assist you.

1. You will be provided with six folders. Each folder is labelled with the school's name and contains 30 copies of the questionnaire. A form asking for the total number of teachers at the schools within the four subject areas is provided in each folder as well.
2. The questionnaire will take less than 10 minutes to complete so I would appreciate if you distribute and collect the same day rather than distribute and then return to collect. Please remind the teachers to not write their name or the name of their schools on the questionnaire. Please ensure that the questionnaires collected from the schools are not mixed up, but placed according to the label on the Manilla folder.
3. **Please try your best to ensure that all math, English, social studies and science teachers at each school complete the entire questionnaire as accurately an honestly as possible**.
4. When you are finished distributing and collecting, please message me on Facebook and then I will let someone collect them to mail to me.

Thanks again for your help. I will discuss payment with you through Facebook.

APPENDIX F: LETTER TO TEACHERS

9 May 2016

Dear Teacher,

I am a graduate student in the School of Education at the University of the South Pacific currently enrolled in the Master of Arts in Education programme. This is a request for your cooperation in a study I am conducting entitled *Secondary Teachers' Assessment Beliefs and Practices*. The purpose of this study is to obtain information on high school teachers' beliefs of assessment and their classroom assessment practices. It will also examine the relationship between assessment beliefs and classroom assessment practices.

Approval to conduct this study has been obtained from the Research Ethics Committee, University of the South Pacific, from Mr. Marvin Kellyman, General Manager of Government Schools and from your school principal.

The attached survey instrument will require approximately 15 minutes to complete and will be used to gather information regarding high school teachers' assessment beliefs and practices. Your completion of the survey implies that you have consented to participate. To ensure anonymity, please do not put your name or your school name on the survey. Return the completed survey to Ms. Trycelin Lord, the researcher assistant.

The promise of strict confidentiality is assured in both the collection and reporting of the data. Any findings obtained in connection with this study will be presented in such a way that no individual or school will be identifiable. All data will be securely maintained by Dr. Ledua Waqailiti, School of Education, University of the South Pacific, for a minimum of five years and then destroyed in accordance with recommended guidelines.

I am hopeful that results from this study may assist universities and districts with preparing and training teachers to utilize assessment practices in ways that enhance instructional planning and student learning. Should you have any questions about this study, please contact me at s11119266@student.usp.ac.fj.

Thank you in advance for your time and willingness to share your assessment beliefs and practices. This study could not be completed without your help.

Sincerely,

____

Gina Middleton
Master's Student
University of the South Pacific

APPENDIX G: PERMISSION TO USE INSTRUMENT

Dear Mr. Gavin Brown:

I am a student in the Master of Arts in Education programme at the University of South Pacific and I am currently writing my master thesis titled *Assessment Conception and Practices of Secondary School Teachers in Belize,* under the direction of my supervisor, Dr. Ledua Waqailiti, who can be reached at 679-323-2296 or ledua.waqailiti@usp.ac.fj.

I am seeking your permission to use the Teachers' Conception of Assessment Inventory (TCoA) in my research study. Specially, I would like permission to use the Greek version used in Brown and Michalis (2011) since, like Cyprus, there is no previous research studies on assessment conception in Belize. I would like to use and print your survey under the following conditions:

- I will use the survey only for my research study and will not sell or use it with any compensated or curriculum development activities.
- I will include the copyright statement on all copies of the instrument.
- I will send a copy of my completed research study to your attention upon completion of the study.

If these are acceptable terms and conditions, please indicate so by replying to me through e-mail: s11119266@student.usp.ac.fj

Sincerely,

Gina Middleton

Master's Degree Candidate

USP
THE UNIVERSITY OF THE SOUTH PACIFIC

Gina Middleton <s11119266@student.usp.ac.fj>

use TCoA in Belize

1 message

Gavin Brown <gt.brown@auckland.ac.nz> 12 April 2016 at 06:01
To: "s11119266@student.usp.ac.fj" <s11119266@student.usp.ac.fj>

dear Gina

thank you for the clarification of what you want

I suspect you don't want the Greek language version of the questionnaire—is Greek spoken in Belize?

I suspect you want to know how to reconstruct the Greek analysis version from the items in the survey

I have attached the Greek language version and the English versions that were compared in the paper

Table 4 of the 2011 Brown & Michaelides paper provides the items in both English and Greek that made up Model D

The correct citation for the TCoA is

Brown, G. T. L. (2006). *Teachers' conceptions of assessment inventory–Abridged* (TCoA-IIIA-Version 3-Abridged). Unpublished test. Auckland, NZ: University of Auckland.

I hope this is what you wanted to know and wish you all the best in your studies

sincerely

Prof. Gavin T L Brown, PhD

Director Quantitative Data Analysis and Research Unit

Faculty of Education & Social Work

The University of Auckland

Tel: +64 9 3737599 ext. 48602

Mob: +64 22 108 7253

New available: Panadero, E., Brown, G. T. L., & Strijbos, J-W. (2015). The future of student self-assessment: A review of known unknowns and potential directions. *Educational Psychology Review*. Advance online publication. doi: 10.1007/s10648-015-9350-2 (http://link.springer.com/article/10.1007%2Fs10648-015-9350-2)

APPENDIX H: SIDAK'S POST HOC TESTS

Table: H.1. *Independent Samples Test for Gender and Assessment Practices*

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Self-Assessment	Equal variances assumed	.073	.788	-.630	56	.532	-.24067	.38232	-1.00654	.52520
	Equal variances not assumed			-.630	41.832	.532	-.24067	.38187	-1.01140	.53007
Individual Project	Equal variances assumed	.029	.865	-.236	56	.814	-.05019	.21240	-.47569	.37530
	Equal variances not assumed			-.240	43.888	.811	-.05019	.20875	-.47092	.37054
Major Exams	Equal variances assumed	.187	.667	-.330	56	.742	-.09266	.28041	-.65439	.46906
	Equal variances not assumed			-.320	37.818	.751	-.09266	.28961	-.67905	.49372
Observation	Equal variances assumed	1.096	.300	-.505	55	.616	-.19054	.37747	-.94701	.56593
	Equal variances not assumed			-.562	51.463	.577	-.19054	.33916	-.87129	.49021
Group Project	Equal variances assumed	.128	.722	-1.038	56	.304	-.21493	.20704	-.62967	.19982
	Equal variances not assumed			-1.043	42.276	.303	-.21493	.20606	-.63070	.20084
Self-Made Test with Select Type Items	Equal variances assumed	2.335	.132	.501	56	.618	.18662	.37239	-.55937	.93260
	Equal variances not assumed			.465	33.301	.645	.18662	.40145	-.62987	1.00310
Oral Presentations	Equal variances assumed	.001	.970	-.504	56	.616	-.13127	.26041	-.65293	.39038
	Equal variances not assumed			-.510	43.263	.612	-.13127	.25717	-.64983	.38728
Provided by Publishers	Equal variances assumed	.161	.690	1.119	56	.268	.32561	.29108	-.25749	.90871
	Equal variances not assumed			1.113	41.047	.272	.32561	.29258	-.26524	.91647
Portfolio	Equal variances assumed	1.479	.229	.255	54	.799	.08571	.33573	-.58738	.75881
	Equal variances not assumed			.271	50.111	.787	.08571	.31576	-.54847	.71990
Peer Assessment	Equal variances assumed	.555	.459	.041	56	.967	.01416	.34359	-.67414	.70246

	Equal variances not assumed			.040	36.847	.969	.01416	.35795	-.71122	.73953
Questioning	Equal variances assumed	2.101	.153	1.983	56	.052	.61776	.31155	-.00635	1.24187
	Equal variances not assumed			1.874	35.165	.069	.61776	.32969	-.05144	1.28696
Authentic Assessment	Equal variances assumed	.835	.365	.692	56	.492	.21750	.31452	-.41255	.84756
	Equal variances not assumed			.676	39.007	.503	.21750	.32152	-.43284	.86784
Self-made Test with Supply-type Items	Equal variances assumed	1.335	.253	1.401	56	.167	.37709	.26908	-.16195	.91613
	Equal variances not assumed			1.472	47.991	.148	.37709	.25617	-.13797	.89215
Essay	Equal variances assumed	.034	.854	-.337	54	.738	-.13229	.39312	-.92044	.65586
	Equal variances not assumed			-.344	38.742	.733	-.13229	.38436	-.90991	.64533
Performance Assessment	Equal variances assumed	4.183	.046	-.063	53	.950	-.02193	.34956	-.72306	.67920
	Equal variances not assumed			-.072	51.050	.943	-.02193	.30544	-.63512	.59126

Table: H.2. *Independent Samples Test for Gender and Usage of Assessment Information*

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Make Decision about Individual Student	Equal variances assumed	.237	.628	-.102	56	.919	-.03089	.30180	-.63547	.57369
	Equal variances not assumed			-.098	36.873	.922	-.03089	.31434	-.66787	.60610
Promotion and Graduation	Equal variances assumed	2.544	.116	1.378	56	.174	.33719	.24462	-.15283	.82722
	Equal variances not assumed			1.252	31.368	.220	.33719	.26929	-.21176	.88615
Plan Teaching	Equal variances assumed	.060	.808	.101	56	.920	.01931	.19180	-.36492	.40353
	Equal variances not assumed			.099	39.774	.922	.01931	.19482	-.37450	.41311
Help Students Address Problems in Learning	Equal variances assumed	.370	.546	-.197	55	.845	-.03968	.20177	-.44404	.36468
	Equal variances not assumed			-.193	39.536	.848	-.03968	.20571	-.45558	.37622

Assign Grade	Equal variances assumed	9.373	.003	1.876	55	.066	.43651	.23272	-.02988	.90289
	Equal variances not assumed			1.603	26.331	.121	.43651	.27228	-.12283	.99585
Provide Descriptive Feedback	Equal variances assumed	.488	.488	1.089	56	.281	.27799	.25530	-.23343	.78941
	Equal variances not assumed			1.003	32.666	.323	.27799	.27706	-.28592	.84190
Rank Student	Equal variances assumed	.740	.393	.707	56	.482	.20978	.29656	-.38430	.80386
	Equal variances not assumed			.643	31.397	.525	.20978	.32636	-.45550	.87506
Assessment without Providing Grade	Equal variances assumed	.305	.583	.772	54	.443	.35000	.45312	-.55845	1.25845
	Equal variances not assumed			.760	37.561	.452	.35000	.46045	-.58249	1.28249

Table: H.3. *Independent Samples Test for Gender and Assessment Conceptions*

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Improvement	Equal variances assumed	.032	.859	1.158	56	.252	.24082	.20793	-.17571	.65734
	Equal variances not assumed			1.130	38.745	.265	.24082	.21303	-.19018	.67181
Student_Accountability	Equal variances assumed	2.962	.091	1.427	56	.159	.35307	.24748	-.14270	.84884
	Equal variances not assumed			1.328	33.640	.193	.35307	.26587	-.18746	.89360
School_Accountability	Equal variances assumed	.046	.831	.475	56	.637	.13792	.29028	-.44357	.71942
	Equal variances not assumed			.467	39.663	.643	.13792	.29511	-.45868	.73452
Irrelevancy	Equal variances assumed	3.128	.082	-1.251	56	.216	-.24556	.19628	-.63875	.14763
	Equal variances not assumed			-1.319	48.455	.193	-.24556	.18617	-.61979	.12867

Table: H.4. *ANOVA Test for Subject Taught and Assessment Practices*

		Sum of Squares	df	Mean Square	F	Sig.
Self-Assessment	Between Groups	1.368	3	.456	.227	.877
	Within Groups	102.378	51	2.007		
	Total	103.745	54			
Individual Project	Between Groups	.850	3	.283	.447	.721
	Within Groups	32.350	51	.634		
	Total	33.200	54			
Major Exams	Between Groups	4.186	3	1.395	1.312	.281
	Within Groups	54.250	51	1.064		
	Total	58.436	54			
Observation	Between Groups	3.869	3	1.290	.675	.571
	Within Groups	95.464	50	1.909		
	Total	99.333	53			
Group Project	Between Groups	1.134	3	.378	.630	.599
	Within Groups	30.611	51	.600		
	Total	31.745	54			
Self-Made Test with Select Type Items	Between Groups	16.468	3	5.489	3.208	.031
	Within Groups	87.278	51	1.711		
	Total	103.745	54			
Oral Presentations	Between Groups	2.866	3	.955	1.062	.373
	Within Groups	45.861	51	.899		
	Total	48.727	54			
Provided by Publishers	Between Groups	4.018	3	1.339	1.352	.268
	Within Groups	50.528	51	.991		
	Total	54.545	54			
Portfolio	Between Groups	2.939	3	.980	.749	.528
	Within Groups	64.080	49	1.308		
	Total	67.019	52			
Peer Assessment	Between Groups	4.033	3	1.344	.953	.422
	Within Groups	71.967	51	1.411		
	Total	76.000	54			
Questioning	Between Groups	3.089	3	1.030	.741	.533
	Within Groups	70.911	51	1.390		
	Total	74.000	54			
Authentic Assessment	Between Groups	3.633	3	1.211	.904	.446
	Within Groups	68.294	51	1.339		
	Total	71.927	54			
Self-made Test with Supply-type Items	Between Groups	4.949	3	1.650	1.664	.186
	Within Groups	50.578	51	.992		
	Total	55.527	54			
Essay	Between Groups	27.199	3	9.066	6.252	.001
	Within Groups	72.505	50	1.450		
	Total	99.704	53			
Performance Assessment	Between Groups	5.540	3	1.847	1.251	.301
	Within Groups	73.793	50	1.476		
	Total	79.333	53			

Table: H.5. *Sidak's Post Hoc Test for Subject Taught and Assessment Practices*

Dependent Variable	(I) Subject	(J) Subject	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Self-Assessment	Math	English	.11111	.49533	1.000	-1.2446	1.4668
		Social Studies	.16667	.54874	1.000	-1.3352	1.6686
		Science	.46667	.57842	.963	-1.1165	2.0498

	English	Math	-.11111	.49533	1.000	-1.4668	1.2446
		Social Studies	.05556	.52802	1.000	-1.3897	1.5008
		Science	.35556	.55881	.989	-1.1739	1.8850
	Social Studies	Math	-.16667	.54874	1.000	-1.6686	1.3352
		English	-.05556	.52802	1.000	-1.5008	1.3897
		Science	.30000	.60665	.997	-1.3604	1.9604
	Science	Math	-.46667	.57842	.963	-2.0498	1.1165
		English	-.35556	.55881	.989	-1.8850	1.1739
		Social Studies	-.30000	.60665	.997	-1.9604	1.3604
Individual Project	Math	English	-.06667	.27844	1.000	-.8288	.6954
		Social Studies	-.31667	.30846	.892	-1.1609	.5276
		Science	-.23333	.32514	.979	-1.1233	.6566
	English	Math	.06667	.27844	1.000	-.6954	.8288
		Social Studies	-.25000	.29681	.955	-1.0624	.5624
		Science	-.16667	.31412	.996	-1.0264	.6931
	Social Studies	Math	.31667	.30846	.892	-.5276	1.1609
		English	.25000	.29681	.955	-.5624	1.0624
		Science	.08333	.34101	1.000	-.8500	1.0167
	Science	Math	.23333	.32514	.979	-.6566	1.1233
		English	.16667	.31412	.996	-.6931	1.0264
		Social Studies	-.08333	.34101	1.000	-1.0167	.8500
Major Exams	Math	English	-.20000	.36057	.995	-1.1869	.7869
		Social Studies	.05000	.39945	1.000	-1.0433	1.1433
		Science	-.73333	.42106	.423	-1.8858	.4191
	English	Math	.20000	.36057	.995	-.7869	1.1869
		Social Studies	.25000	.38437	.988	-.8020	1.3020
		Science	-.53333	.40678	.729	-1.6467	.5800
	Social Studies	Math	-.05000	.39945	1.000	-1.1433	1.0433
		English	-.25000	.38437	.988	-1.3020	.8020
		Science	-.78333	.44161	.402	-1.9920	.4254
	Science	Math	.73333	.42106	.423	-.4191	1.8858
		English	.53333	.40678	.729	-.5800	1.6467
		Social Studies	.78333	.44161	.402	-.4254	1.9920
Observation	Math	English	.02381	.49239	1.000	-1.3250	1.3726
		Social Studies	-.39286	.54359	.979	-1.8819	1.0962
		Science	-.64286	.57211	.844	-2.2100	.9243
	English	Math	-.02381	.49239	1.000	-1.3726	1.3250
		Social Studies	-.41667	.51495	.963	-1.8273	.9939
		Science	-.66667	.54498	.787	-2.1595	.8262
	Social Studies	Math	.39286	.54359	.979	-1.0962	1.8819
		English	.41667	.51495	.963	-.9939	1.8273
		Science	-.25000	.59164	.999	-1.8706	1.3706
	Science	Math	.64286	.57211	.844	-.9243	2.2100
		English	.66667	.54498	.787	-.8262	2.1595
		Social Studies	.25000	.59164	.999	-1.3706	1.8706
Group Project	Math	English	.04444	.27085	1.000	-.6969	.7858
		Social Studies	-.23333	.30005	.969	-1.0546	.5879
		Science	-.30000	.31629	.923	-1.1657	.5657
	English	Math	-.04444	.27085	1.000	-.7858	.6969
		Social Studies	-.27778	.28873	.918	-1.0680	.5125
		Science	-.34444	.30556	.842	-1.1808	.4919
	Social Studies	Math	.23333	.30005	.969	-.5879	1.0546
		English	.27778	.28873	.918	-.5125	1.0680
		Science	-.06667	.33172	1.000	-.9746	.8413
	Science	Math	.30000	.31629	.923	-.5657	1.1657
		English	.34444	.30556	.842	-.4919	1.1808
		Social Studies	.06667	.33172	1.000	-.8413	.9746

Self-Made Test with Select Type Items	Math	English	.21111	.45734	.998	-1.0407	1.4629
		Social Studies	-.23333	.50666	.998	-1.6201	1.1534
		Science	-1.33333	.53406	.091	-2.7951	.1284
	English	Math	-.21111	.45734	.998	-1.4629	1.0407
		Social Studies	-.44444	.48753	.935	-1.7788	.8899
		Science	-1.54444*	.51595	.025	-2.9566	-.1323
	Social Studies	Math	.23333	.50666	.998	-1.1534	1.6201
		English	.44444	.48753	.935	-.8899	1.7788
		Science	-1.10000	.56013	.288	-2.6331	.4331
	Science	Math	1.33333	.53406	.091	-.1284	2.7951
		English	1.54444*	.51595	.025	.1323	2.9566
		Social Studies	1.10000	.56013	.288	-.4331	2.6331
Oral Presentations	Math	English	-.55556	.33152	.468	-1.4629	.3518
		Social Studies	-.41667	.36727	.838	-1.4219	.5886
		Science	-.50000	.38713	.742	-1.5596	.5596
	English	Math	.55556	.33152	.468	-.3518	1.4629
		Social Studies	.13889	.35340	.999	-.8284	1.1062
		Science	.05556	.37401	1.000	-.9681	1.0792
	Social Studies	Math	.41667	.36727	.838	-.5886	1.4219
		English	-.13889	.35340	.999	-1.1062	.8284
		Science	-.08333	.40603	1.000	-1.1947	1.0280
	Science	Math	.50000	.38713	.742	-.5596	1.5596
		English	-.05556	.37401	1.000	-1.0792	.9681
		Social Studies	.08333	.40603	1.000	-1.0280	1.1947
Provided by Publishers	Math	English	-.07778	.34798	1.000	-1.0302	.8747
		Social Studies	.61667	.38550	.522	-.4385	1.6718
		Science	.00000	.40635	1.000	-1.1122	1.1122
	English	Math	.07778	.34798	1.000	-.8747	1.0302
		Social Studies	.69444	.37095	.340	-.3209	1.7097
		Science	.07778	.39258	1.000	-.9967	1.1523
	Social Studies	Math	-.61667	.38550	.522	-1.6718	.4385
		English	-.69444	.37095	.340	-1.7097	.3209
		Science	-.61667	.42619	.633	-1.7832	.5498
	Science	Math	.00000	.40635	1.000	-1.1122	1.1122
		English	-.07778	.39258	1.000	-1.1523	.9967
		Social Studies	.61667	.42619	.633	-.5498	1.7832
Portfolio	Math	English	-.34902	.40511	.950	-1.4596	.7616
		Social Studies	.26061	.45395	.994	-.9839	1.5051
		Science	.13333	.46686	1.000	-1.1466	1.4132
	English	Math	.34902	.40511	.950	-.7616	1.4596
		Social Studies	.60963	.44251	.684	-.6035	1.8228
		Science	.48235	.45574	.877	-.7671	1.7318
	Social Studies	Math	-.26061	.45395	.994	-1.5051	.9839
		English	-.60963	.44251	.684	-1.8228	.6035
		Science	-.12727	.49966	1.000	-1.4971	1.2426
	Science	Math	-.13333	.46686	1.000	-1.4132	1.1466
		English	-.48235	.45574	.877	-1.7318	.7671
		Social Studies	.12727	.49966	1.000	-1.2426	1.4971
Peer Assessment	Math	English	-.53333	.41529	.747	-1.6700	.6033
		Social Studies	.13333	.46007	1.000	-1.1259	1.3926
		Science	-.30000	.48496	.990	-1.6274	1.0274
	English	Math	.53333	.41529	.747	-.6033	1.6700
		Social Studies	.66667	.44270	.591	-.5450	1.8784
		Science	.23333	.46851	.997	-1.0490	1.5157
	Social Studies	Math	-.13333	.46007	1.000	-1.3926	1.1259
		English	-.66667	.44270	.591	-1.8784	.5450
		Science	-.43333	.50863	.953	-1.8255	.9588

	Science	Math	.30000	.48496	.990	-1.0274	1.6274
		English	-.23333	.46851	.997	-1.5157	1.0490
		Social Studies	.43333	.50863	.953	-.9588	1.8255
Questioning	Math	English	-.41111	.41224	.904	-1.5394	.7172
		Social Studies	.20000	.45669	.999	-1.0500	1.4500
		Science	-.23333	.48139	.997	-1.5509	1.0843
	English	Math	.41111	.41224	.904	-.7172	1.5394
		Social Studies	.61111	.43945	.674	-.5917	1.8139
		Science	.17778	.46507	.999	-1.0951	1.4507
	Social Studies	Math	-.20000	.45669	.999	-1.4500	1.0500
		English	-.61111	.43945	.674	-1.8139	.5917
		Science	-.43333	.50489	.951	-1.8152	.9486
	Science	Math	.23333	.48139	.997	-1.0843	1.5509
		English	-.17778	.46507	.999	-1.4507	1.0951
		Social Studies	.43333	.50489	.951	-.9486	1.8152
Authentic Assessment	Math	English	-.55556	.40456	.686	-1.6629	.5517
		Social Studies	-.08333	.44818	1.000	-1.3100	1.1434
		Science	-.53333	.47242	.841	-1.8264	.7597
	English	Math	.55556	.40456	.686	-.5517	1.6629
		Social Studies	.47222	.43126	.859	-.7082	1.6526
		Science	.02222	.45640	1.000	-1.2270	1.2714
	Social Studies	Math	.08333	.44818	1.000	-1.1434	1.3100
		English	-.47222	.43126	.859	-1.6526	.7082
		Science	-.45000	.49548	.936	-1.8062	.9062
	Science	Math	.53333	.47242	.841	-.7597	1.8264
		English	-.02222	.45640	1.000	-1.2714	1.2270
		Social Studies	.45000	.49548	.936	-.9062	1.8062
Self-made Test with Supply-type Items	Math	English	-.47778	.34815	.687	-1.4307	.4751
		Social Studies	-.03333	.38569	1.000	-1.0890	1.0223
		Science	-.76667	.40655	.332	-1.8794	.3461
	English	Math	.47778	.34815	.687	-.4751	1.4307
		Social Studies	.44444	.37113	.802	-.5714	1.4602
		Science	-.28889	.39277	.977	-1.3639	.7861
	Social Studies	Math	.03333	.38569	1.000	-1.0223	1.0890
		English	-.44444	.37113	.802	-1.4602	.5714
		Science	-.73333	.42640	.438	-1.9004	.4337
	Science	Math	.76667	.40655	.332	-.3461	1.8794
		English	.28889	.39277	.977	-.7861	1.3639
		Social Studies	.73333	.42640	.438	-.4337	1.9004
Essay	Math	English	-1.51111*	.42099	.005	-2.6643	-.3579
		Social Studies	-1.61212*	.47802	.009	-2.9215	-.3027
		Science	-1.66667*	.49161	.008	-3.0133	-.3200
	English	Math	1.51111*	.42099	.005	.3579	2.6643
		Social Studies	-.10101	.46086	1.000	-1.3634	1.1614
		Science	-.15556	.47494	1.000	-1.4565	1.1454
	Social Studies	Math	1.61212*	.47802	.009	.3027	2.9215
		English	.10101	.46086	1.000	-1.1614	1.3634
		Science	-.05455	.52615	1.000	-1.4958	1.3867
	Science	Math	1.66667*	.49161	.008	.3200	3.0133
		English	.15556	.47494	1.000	-1.1454	1.4565
		Social Studies	.05455	.52615	1.000	-1.3867	1.4958
Performance Assessment	Math	English	-.67778	.42471	.525	-1.8412	.4856
		Social Studies	-.67273	.48224	.671	-1.9937	.6483
		Science	-.80000	.49596	.513	-2.1586	.5586
	English	Math	.67778	.42471	.525	-.4856	1.8412
		Social Studies	.00505	.46493	1.000	-1.2685	1.2786

	Science	-.12222	.47914	1.000	-1.4347	1.1903
Social Studies	Math	.67273	.48224	.671	-.6483	1.9937
	English	-.00505	.46493	1.000	-1.2786	1.2685
	Science	-.12727	.53081	1.000	-1.5813	1.3267
Science	Math	.80000	.49596	.513	-.5586	2.1586
	English	.12222	.47914	1.000	-1.1903	1.4347
	Social Studies	.12727	.53081	1.000	-1.3267	1.5813

*. The mean difference is significant at the 0.05 level.

Table: H.6. *ANOVA Test for Subject Taught and Usage of Assessment Information*

		Sum of Squares	df	Mean Square	F	Sig.
Make Decision about Individual Student	Between Groups	1.704	3	.568	.455	.715
	Within Groups	63.678	51	1.249		
	Total	65.382	54			
Promotion and Graduation	Between Groups	2.025	3	.675	.896	.450
	Within Groups	38.411	51	.753		
	Total	40.436	54			
Plan Teaching	Between Groups	4.037	3	1.346	3.099	.035
	Within Groups	22.144	51	.434		
	Total	26.182	54			
Help Students Address Problems in Learning	Between Groups	7.462	3	2.487	5.901	.002
	Within Groups	21.075	50	.422		
	Total	28.537	53			
Assign Grade	Between Groups	6.880	3	2.293	3.424	.024
	Within Groups	33.490	50	.670		
	Total	40.370	53			
Provide Descriptive Feedback	Between Groups	1.918	3	.639	1.037	.384
	Within Groups	31.428	51	.616		
	Total	33.345	54			
Rank Student	Between Groups	3.933	3	1.311	1.466	.235
	Within Groups	45.594	51	.894		
	Total	49.527	54			
Assessment without Providing Grade	Between Groups	9.976	3	3.325	1.272	.294
	Within Groups	128.099	49	2.614		
	Total	138.075	52			

Table: H.7. *Sidak's Post Hoc Test for Subject Taught and Usage of Assessment Information*

Dependent Variable	(I) Subject	(J) Subject	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Make Decision about Individual Student	Math	English	.42222	.39065	.866	-.6470	1.4914
		Social Studies	.36667	.43277	.954	-.8178	1.5512
		Science	.36667	.45618	.964	-.8819	1.6152
	English	Math	-.42222	.39065	.866	-1.4914	.6470
		Social Studies	-.05556	.41643	1.000	-1.1953	1.0842
		Science	-.05556	.44071	1.000	-1.2618	1.1507
	Social Studies	Math	-.36667	.43277	.954	-1.5512	.8178
		English	.05556	.41643	1.000	-1.0842	1.1953
		Science	.00000	.47844	1.000	-1.3095	1.3095
	Science	Math	-.36667	.45618	.964	-1.6152	.8819
		English	.05556	.44071	1.000	-1.1507	1.2618
		Social Studies	.00000	.47844	1.000	-1.3095	1.3095
Promotion and Graduation	Math	English	.14444	.30340	.998	-.6860	.9749
		Social Studies	.53333	.33612	.532	-.3866	1.4533
		Science	.13333	.35430	.999	-.8364	1.1031

	English	Math	-.14444	.30340	.998	-.9749	.6860
		Social Studies	.38889	.32343	.799	-.4963	1.2741
		Science	-.01111	.34228	1.000	-.9480	.9257
	Social Studies	Math	-.53333	.33612	.532	-1.4533	.3866
		English	-.38889	.32343	.799	-1.2741	.4963
		Science	-.40000	.37159	.868	-1.4171	.6171
	Science	Math	-.13333	.35430	.999	-1.1031	.8364
		English	.01111	.34228	1.000	-.9257	.9480
		Social Studies	.40000	.37159	.868	-.6171	1.4171
Plan Teaching	Math	English	.65556*	.23037	.038	.0250	1.2861
		Social Studies	.60000	.25521	.128	-.0985	1.2985
		Science	.40000	.26901	.604	-.3363	1.1363
	English	Math	-.65556*	.23037	.038	-1.2861	-.0250
		Social Studies	-.05556	.24557	1.000	-.7277	.6166
		Science	-.25556	.25989	.910	-.9669	.4558
	Social Studies	Math	-.60000	.25521	.128	-1.2985	.0985
		English	.05556	.24557	1.000	-.6166	.7277
		Science	-.20000	.28214	.981	-.9722	.5722
	Science	Math	-.40000	.26901	.604	-1.1363	.3363
		English	.25556	.25989	.910	-.4558	.9669
		Social Studies	.20000	.28214	.981	-.5722	.9722
Help Students Address Problems in Learning	Math	English	.58730	.23135	.083	-.0464	1.2210
		Social Studies	1.05952*	.25541	.001	.3599	1.7591
		Science	.64286	.26881	.117	-.0935	1.3792
	English	Math	-.58730	.23135	.083	-1.2210	.0464
		Social Studies	.47222	.24196	.295	-.1906	1.1350
		Science	.05556	.25606	1.000	-.6459	.7570
	Social Studies	Math	-1.05952*	.25541	.001	-1.7591	-.3599
		English	-.47222	.24196	.295	-1.1350	.1906
		Science	-.41667	.27799	.596	-1.1781	.3448
	Science	Math	-.64286	.26881	.117	-1.3792	.0935
		English	-.05556	.25606	1.000	-.7570	.6459
		Social Studies	.41667	.27799	.596	-.3448	1.1781
Assign Grade	Math	English	.15556	.28612	.995	-.6282	.9393
		Social Studies	.96364*	.32488	.027	.0737	1.8535
		Science	.50000	.33412	.598	-.4152	1.4152
	English	Math	-.15556	.28612	.995	-.9393	.6282
		Social Studies	.80808	.31321	.075	-.0499	1.6660
		Science	.34444	.32279	.873	-.5397	1.2286
	Social Studies	Math	-.96364*	.32488	.027	-1.8535	-.0737
		English	-.80808	.31321	.075	-1.6660	.0499
		Science	-.46364	.35759	.739	-1.4432	.5159
	Science	Math	-.50000	.33412	.598	-1.4152	.4152
		English	-.34444	.32279	.873	-1.2286	.5397
		Social Studies	.46364	.35759	.739	-.5159	1.4432
Provide Descriptive Feedback	Math	English	.38889	.27444	.655	-.3623	1.1400
		Social Studies	.08333	.30403	1.000	-.7488	.9155
		Science	.43333	.32048	.701	-.4438	1.3105
	English	Math	-.38889	.27444	.655	-1.1400	.3623
		Social Studies	-.30556	.29255	.884	-1.1063	.4952
		Science	.04444	.30961	1.000	-.8030	.8919
	Social Studies	Math	-.08333	.30403	1.000	-.9155	.7488
		English	.30556	.29255	.884	-.4952	1.1063
		Science	.35000	.33612	.885	-.5700	1.2700
	Science	Math	-.43333	.32048	.701	-1.3105	.4438
		English	-.04444	.30961	1.000	-.8919	.8030
		Social Studies	-.35000	.33612	.885	-1.2700	.5700

Rank Student	Math	English	.18889	.33056	.994	-.7159	1.0936
		Social Studies	.71667	.36620	.292	-.2856	1.7190
		Science	.46667	.38601	.795	-.5899	1.5232
	English	Math	-.18889	.33056	.994	-1.0936	.7159
		Social Studies	.52778	.35237	.596	-.4367	1.4922
		Science	.27778	.37292	.975	-.7429	1.2985
	Social Studies	Math	-.71667	.36620	.292	-1.7190	.2856
		English	-.52778	.35237	.596	-1.4922	.4367
		Science	-.25000	.40485	.990	-1.3581	.8581
	Science	Math	-.46667	.38601	.795	-1.5232	.5899
		English	-.27778	.37292	.975	-1.2985	.7429
		Social Studies	.25000	.40485	.990	-.8581	1.3581
Assessment without Providing Grade	Math	English	-1.03922	.57277	.377	-2.6095	.5311
		Social Studies	-.15152	.64183	1.000	-1.9111	1.6081
		Science	-.53333	.66008	.963	-2.3430	1.2763
	English	Math	1.03922	.57277	.377	-.5311	2.6095
		Social Studies	.88770	.62565	.654	-.8275	2.6030
		Science	.50588	.64437	.968	-1.2607	2.2724
	Social Studies	Math	.15152	.64183	1.000	-1.6081	1.9111
		English	-.88770	.62565	.654	-2.6030	.8275
		Science	-.38182	.70646	.995	-2.3186	1.5550
	Science	Math	.53333	.66008	.963	-1.2763	2.3430
		English	-.50588	.64437	.968	-2.2724	1.2607
		Social Studies	.38182	.70646	.995	-1.5550	2.3186

*. The mean difference is significant at the 0.05 level.

Table: H.8. *ANOVA Test for Subject Taught and Assessment Conceptions*

		Sum of Squares	df	Mean Square	F	Sig.
Improvement	Between Groups	3.424	3	1.141	2.003	.125
	Within Groups	29.061	51	.570		
	Total	32.485	54			
Student_Accountability	Between Groups	2.697	3	.899	1.051	.378
	Within Groups	43.619	51	.855		
	Total	46.315	54			
School_Accountability	Between Groups	5.800	3	1.933	1.828	.154
	Within Groups	53.946	51	1.058		
	Total	59.745	54			
Irrelevancy	Between Groups	.754	3	.251	.446	.721
	Within Groups	28.714	51	.563		
	Total	29.468	54			

Table: H.9. *ANOVA Test for Teaching Experiences and Assessment Practices*

		Sum of Squares	df	Mean Square	F	Sig.
Self-Assessment	Between Groups	4.656	3	1.552	.780	.511
	Within Groups	105.485	53	1.990		
	Total	110.140	56			
Individual Project	Between Groups	1.032	3	.344	.561	.643
	Within Groups	32.477	53	.613		
	Total	33.509	56			
Major Exams	Between Groups	3.949	3	1.316	1.268	.295
	Within Groups	55.033	53	1.038		
	Total	58.982	56			
Observation	Between Groups	5.185	3	1.728	.929	.434
	Within Groups	96.797	52	1.861		
	Total	101.982	55			
Group Project	Between Groups	.742	3	.247	.415	.743

	Within Groups	31.574	53	.596		
	Total	32.316	56			
Self-Made Test with Select Type Items	Between Groups	7.530	3	2.510	1.376	.260
	Within Groups	96.715	53	1.825		
	Total	104.246	56			
Oral Presentations	Between Groups	2.030	3	.677	.733	.537
	Within Groups	48.952	53	.924		
	Total	50.982	56			
Provided by Publishers	Between Groups	7.701	3	2.567	2.375	.080
	Within Groups	57.282	53	1.081		
	Total	64.982	56			
Portfolio	Between Groups	.899	3	.300	.194	.900
	Within Groups	78.810	51	1.545		
	Total	79.709	54			
Peer Assessment	Between Groups	6.858	3	2.286	1.483	.230
	Within Groups	81.703	53	1.542		
	Total	88.561	56			
Questioning	Between Groups	2.824	3	.941	.664	.578
	Within Groups	75.106	53	1.417		
	Total	77.930	56			
Authentic Assessment	Between Groups	2.614	3	.871	.649	.587
	Within Groups	71.106	53	1.342		
	Total	73.719	56			
Self-made Test with Supply-type Items	Between Groups	2.540	3	.847	.839	.479
	Within Groups	53.495	53	1.009		
	Total	56.035	56			
Essay	Between Groups	2.589	3	.863	.438	.727
	Within Groups	102.393	52	1.969		
	Total	104.982	55			
Performance Assessment	Between Groups	12.753	3	4.251	3.198	.031
	Within Groups	67.793	51	1.329		
	Total	80.545	54			

Table: H.10. *Sidak's Post Hoc Test for Teaching Experiences and Assessment Practices*

Dependent Variable	(I) Teaching Experience	(J) Teaching Experience	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Self-Assessment	Less than 3 years	Between 4 - 10 Years	-.50877	.57087	.941	-2.0689	1.0514
		Between 11 - 20 Years	.04167	.55143	1.000	-1.4653	1.5487
		More than 20 Years	.33333	.78689	.999	-1.8172	2.4838
	Between 4 - 10 Years	Less than 3 years	.50877	.57087	.941	-1.0514	2.0689
		Between 11 - 20 Years	.55044	.43322	.756	-.6335	1.7344
		More than 20 Years	.84211	.70909	.808	-1.0958	2.7800
	Between 11 - 20 Years	Less than 3 years	-.04167	.55143	1.000	-1.5487	1.4653
		Between 4 - 10 Years	-.55044	.43322	.756	-1.7344	.6335
		More than 20 Years	.29167	.69353	.999	-1.6037	2.1870
	More than 20 Years	Less than 3 years	-.33333	.78689	.999	-2.4838	1.8172
		Between 4 - 10 Years	-.84211	.70909	.808	-2.7800	1.0958
		Between 11 - 20 Years	-.29167	.69353	.999	-2.1870	1.6037
Individual Project	Less than 3 years	Between 4 - 10 Years	.18713	.31676	.992	-.6785	1.0528
		Between 11 - 20 Years	.13889	.30597	.998	-.6973	.9751
		More than 20 Years	.55556	.43662	.755	-.6377	1.7488
	Between 4 - 10 Years	Less than 3 years	-.18713	.31676	.992	-1.0528	.6785
		Between 11 - 20 Years	-.04825	.24038	1.000	-.7052	.6087
		More than 20 Years	.36842	.39345	.927	-.7068	1.4437
		Less than 3 years	-.13889	.30597	.998	-.9751	.6973

	Between 11 - 20 Years	Between 4 - 10 Years	.04825	.24038	1.000	-.6087	.7052
		More than 20 Years	.41667	.38482	.865	-.6350	1.4683
	More than 20 Years	Less than 3 years	-.55556	.43662	.755	-1.7488	.6377
		Between 4 - 10 Years	-.36842	.39345	.927	-1.4437	.7068
		Between 11 - 20 Years	-.41667	.38482	.865	-1.4683	.6350
Major Exams	Less than 3 years	Between 4 - 10 Years	-.33333	.41234	.963	-1.4602	.7936
		Between 11 - 20 Years	.25000	.39830	.990	-.8385	1.3385
		More than 20 Years	.26667	.56837	.998	-1.2866	1.8200
	Between 4 - 10 Years	Less than 3 years	.33333	.41234	.963	-.7936	1.4602
		Between 11 - 20 Years	.58333	.31292	.344	-.2718	1.4385
		More than 20 Years	.60000	.51218	.817	-.7997	1.9997
	Between 11 - 20 Years	Less than 3 years	-.25000	.39830	.990	-1.3385	.8385
		Between 4 - 10 Years	-.58333	.31292	.344	-1.4385	.2718
		More than 20 Years	.01667	.50094	1.000	-1.3523	1.3857
	More than 20 Years	Less than 3 years	-.26667	.56837	.998	-1.8200	1.2866
		Between 4 - 10 Years	-.60000	.51218	.817	-1.9997	.7997
		Between 11 - 20 Years	-.01667	.50094	1.000	-1.3857	1.3523
Observation	Less than 3 years	Between 4 - 10 Years	-.90058	.55209	.499	-2.4105	.6093
		Between 11 - 20 Years	-.61111	.53328	.832	-2.0696	.8474
		More than 20 Years	-.86111	.81988	.881	-3.1034	1.3812
	Between 4 - 10 Years	Less than 3 years	.90058	.55209	.499	-.6093	2.4105
		Between 11 - 20 Years	.28947	.41897	.983	-.8564	1.4353
		More than 20 Years	.03947	.75056	1.000	-2.0133	2.0922
	Between 11 - 20 Years	Less than 3 years	.61111	.53328	.832	-.8474	2.0696
		Between 4 - 10 Years	-.28947	.41897	.983	-1.4353	.8564
		More than 20 Years	-.25000	.73684	1.000	-2.2652	1.7652
	More than 20 Years	Less than 3 years	.86111	.81988	.881	-1.3812	3.1034
		Between 4 - 10 Years	-.03947	.75056	1.000	-2.0922	2.0133
		Between 11 - 20 Years	.25000	.73684	1.000	-1.7652	2.2652
Group Project	Less than 3 years	Between 4 - 10 Years	-.28655	.31232	.933	-1.1401	.5670
		Between 11 - 20 Years	-.06944	.30169	1.000	-.8939	.7550
		More than 20 Years	-.04444	.43051	1.000	-1.2210	1.1321
	Between 4 - 10 Years	Less than 3 years	.28655	.31232	.933	-.5670	1.1401
		Between 11 - 20 Years	.21711	.23701	.934	-.4306	.8648
		More than 20 Years	.24211	.38794	.990	-.8181	1.3023
	Between 11 - 20 Years	Less than 3 years	.06944	.30169	1.000	-.7550	.8939
		Between 4 - 10 Years	-.21711	.23701	.934	-.8648	.4306
		More than 20 Years	.02500	.37943	1.000	-1.0119	1.0619
	More than 20 Years	Less than 3 years	.04444	.43051	1.000	-1.1321	1.2210
		Between 4 - 10 Years	-.24211	.38794	.990	-1.3023	.8181
		Between 11 - 20 Years	-.02500	.37943	1.000	-1.0619	1.0119
Self-Made Test with Select Type Items	Less than 3 years	Between 4 - 10 Years	-1.06433	.54663	.296	-2.5582	.4295
		Between 11 - 20 Years	-.63889	.52801	.794	-2.0819	.8041
		More than 20 Years	-1.02222	.75347	.697	-3.0814	1.0369
	Between 4 - 10 Years	Less than 3 years	1.06433	.54663	.296	-.4295	2.5582
		Between 11 - 20 Years	.42544	.41482	.892	-.7082	1.5591
		More than 20 Years	.04211	.67897	1.000	-1.8135	1.8977
		Less than 3 years	.63889	.52801	.794	-.8041	2.0819

	Between 11 - 20 Years	Between 4 - 10 Years	-.42544	.41482	.892	-1.5591	.7082
		More than 20 Years	-.38333	.66408	.993	-2.1982	1.4315
	More than 20 Years	Less than 3 years	1.02222	.75347	.697	-1.0369	3.0814
		Between 4 - 10 Years	-.04211	.67897	1.000	-1.8977	1.8135
		Between 11 - 20 Years	.38333	.66408	.993	-1.4315	2.1982
Oral Presentations	Less than 3 years	Between 4 - 10 Years	-.20468	.38889	.996	-1.2675	.8581
		Between 11 - 20 Years	-.43056	.37565	.832	-1.4572	.5960
		More than 20 Years	.11111	.53605	1.000	-1.3539	1.5761
	Between 4 - 10 Years	Less than 3 years	.20468	.38889	.996	-.8581	1.2675
		Between 11 - 20 Years	-.22588	.29512	.972	-1.0324	.5807
		More than 20 Years	.31579	.48305	.987	-1.0043	1.6359
	Between 11 - 20 Years	Less than 3 years	.43056	.37565	.832	-.5960	1.4572
		Between 4 - 10 Years	.22588	.29512	.972	-.5807	1.0324
		More than 20 Years	.54167	.47245	.831	-.7495	1.8328
	More than 20 Years	Less than 3 years	-.11111	.53605	1.000	-1.5761	1.3539
		Between 4 - 10 Years	-.31579	.48305	.987	-1.6359	1.0043
		Between 11 - 20 Years	-.54167	.47245	.831	-1.8328	.7495
Provided by Publishers	Less than 3 years	Between 4 - 10 Years	.93567	.42068	.169	-.2140	2.0853
		Between 11 - 20 Years	.77778	.40635	.315	-.3327	1.8883
		More than 20 Years	1.37778	.57987	.120	-.2069	2.9625
	Between 4 - 10 Years	Less than 3 years	-.93567	.42068	.169	-2.0853	.2140
		Between 11 - 20 Years	-.15789	.31924	.997	-1.0304	.7146
		More than 20 Years	.44211	.52253	.954	-.9859	1.8701
	Between 11 - 20 Years	Less than 3 years	-.77778	.40635	.315	-1.8883	.3327
		Between 4 - 10 Years	.15789	.31924	.997	-.7146	1.0304
		More than 20 Years	.60000	.51107	.816	-.7967	1.9967
	More than 20 Years	Less than 3 years	-1.37778	.57987	.120	-2.9625	.2069
		Between 4 - 10 Years	-.44211	.52253	.954	-1.8701	.9859
		Between 11 - 20 Years	-.60000	.51107	.816	-1.9967	.7967
Portfolio	Less than 3 years	Between 4 - 10 Years	.32895	.52392	.990	-1.1050	1.7629
		Between 11 - 20 Years	.35870	.51024	.981	-1.0379	1.7553
		More than 20 Years	.15000	.70867	1.000	-1.7897	2.0897
	Between 4 - 10 Years	Less than 3 years	-.32895	.52392	.990	-1.7629	1.1050
		Between 11 - 20 Years	.02975	.38538	1.000	-1.0251	1.0846
		More than 20 Years	-.17895	.62481	1.000	-1.8891	1.5312
	Between 11 - 20 Years	Less than 3 years	-.35870	.51024	.981	-1.7553	1.0379
		Between 4 - 10 Years	-.02975	.38538	1.000	-1.0846	1.0251
		More than 20 Years	-.20870	.61339	1.000	-1.8876	1.4702
	More than 20 Years	Less than 3 years	-.15000	.70867	1.000	-2.0897	1.7897
		Between 4 - 10 Years	.17895	.62481	1.000	-1.5312	1.8891
		Between 11 - 20 Years	.20870	.61339	1.000	-1.4702	1.8876
Peer Assessment	Less than 3 years	Between 4 - 10 Years	.50292	.50241	.902	-.8701	1.8760
		Between 11 - 20 Years	.88889	.48530	.364	-.4374	2.2152
		More than 20 Years	1.15556	.69253	.472	-.7371	3.0482
	Between 4 - 10 Years	Less than 3 years	-.50292	.50241	.902	-1.8760	.8701
		Between 11 - 20 Years	.38596	.38127	.898	-.6560	1.4279
		More than 20 Years	.65263	.62406	.883	-1.0529	2.3581
		Less than 3 years	-.88889	.48530	.364	-2.2152	.4374

	Between 11 - 20 Years	Between 4 - 10 Years	-.38596	.38127	.898	-1.4279	.6560
		More than 20 Years	.26667	.61036	.999	-1.4014	1.9347
	More than 20 Years	Less than 3 years	-1.15556	.69253	.472	-3.0482	.7371
		Between 4 - 10 Years	-.65263	.62406	.883	-2.3581	1.0529
		Between 11 - 20 Years	-.26667	.61036	.999	-1.9347	1.4014
Questionning	Less than 3 years	Between 4 - 10 Years	-.28070	.48170	.993	-1.5971	1.0357
		Between 11 - 20 Years	-.29167	.46530	.990	-1.5633	.9799
		More than 20 Years	-.93333	.66398	.663	-2.7479	.8813
	Between 4 - 10 Years	Less than 3 years	.28070	.48170	.993	-1.0357	1.5971
		Between 11 - 20 Years	-.01096	.36555	1.000	-1.0100	.9881
		More than 20 Years	-.65263	.59833	.861	-2.2878	.9826
	Between 11 - 20 Years	Less than 3 years	.29167	.46530	.990	-.9799	1.5633
		Between 4 - 10 Years	.01096	.36555	1.000	-.9881	1.0100
		More than 20 Years	-.64167	.58520	.858	-2.2410	.9576
	More than 20 Years	Less than 3 years	.93333	.66398	.663	-.8813	2.7479
		Between 4 - 10 Years	.65263	.59833	.861	-.9826	2.2878
		Between 11 - 20 Years	.64167	.58520	.858	-.9576	2.2410
Authentic Assessment	Less than 3 years	Between 4 - 10 Years	-.05263	.46870	1.000	-1.3335	1.2283
		Between 11 - 20 Years	.29167	.45274	.988	-.9456	1.5289
		More than 20 Years	-.40000	.64606	.990	-2.1656	1.3656
	Between 4 - 10 Years	Less than 3 years	.05263	.46870	1.000	-1.2283	1.3335
		Between 11 - 20 Years	.34430	.35569	.915	-.6278	1.3164
		More than 20 Years	-.34737	.58218	.992	-1.9384	1.2437
	Between 11 - 20 Years	Less than 3 years	-.29167	.45274	.988	-1.5289	.9456
		Between 4 - 10 Years	-.34430	.35569	.915	-1.3164	.6278
		More than 20 Years	-.69167	.56941	.791	-2.2478	.8645
	More than 20 Years	Less than 3 years	.40000	.64606	.990	-1.3656	2.1656
		Between 4 - 10 Years	.34737	.58218	.992	-1.2437	1.9384
		Between 11 - 20 Years	.69167	.56941	.791	-.8645	2.2478
Self-made Test with Supply-type Items	Less than 3 years	Between 4 - 10 Years	.14035	.40654	1.000	-.9707	1.2514
		Between 11 - 20 Years	.45833	.39269	.820	-.6148	1.5315
		More than 20 Years	-.13333	.56037	1.000	-1.6648	1.3981
	Between 4 - 10 Years	Less than 3 years	-.14035	.40654	1.000	-1.2514	.9707
		Between 11 - 20 Years	.31798	.30851	.890	-.5251	1.1611
		More than 20 Years	-.27368	.50497	.995	-1.6537	1.1063
	Between 11 - 20 Years	Less than 3 years	-.45833	.39269	.820	-1.5315	.6148
		Between 4 - 10 Years	-.31798	.30851	.890	-1.1611	.5251
		More than 20 Years	-.59167	.49389	.802	-1.9414	.7581
	More than 20 Years	Less than 3 years	.13333	.56037	1.000	-1.3981	1.6648
		Between 4 - 10 Years	.27368	.50497	.995	-1.1063	1.6537
		Between 11 - 20 Years	.59167	.49389	.802	-.7581	1.9414
Essay	Less than 3 years	Between 4 - 10 Years	.21053	.56782	.999	-1.3424	1.7635
		Between 11 - 20 Years	-.26087	.55172	.998	-1.7698	1.2481
		More than 20 Years	.20000	.78269	1.000	-1.9406	2.3406
	Between 4 - 10 Years	Less than 3 years	-.21053	.56782	.999	-1.7635	1.3424
		Between 11 - 20 Years	-.47140	.43503	.865	-1.6612	.7184
		More than 20 Years	-.01053	.70530	1.000	-1.9395	1.9184
		Less than 3 years	.26087	.55172	.998	-1.2481	1.7698

	Between 11 - 20 Years	Between 4 - 10 Years	.47140	.43503	.865	-.7184	1.6612
		More than 20 Years	.46087	.69241	.986	-1.4328	2.3546
	More than 20 Years	Less than 3 years	-.20000	.78269	1.000	-2.3406	1.9406
		Between 4 - 10 Years	.01053	.70530	1.000	-1.9184	1.9395
		Between 11 - 20 Years	-.46087	.69241	.986	-2.3546	1.4328
Performance Assessment	Less than 3 years	Between 4 - 10 Years	-1.27778	.47069	.053	-2.5661	.0105
		Between 11 - 20 Years	-.97585	.45331	.198	-2.2166	.2649
		More than 20 Years	-1.68889	.64308	.066	-3.4490	.0712
	Between 4 - 10 Years	Less than 3 years	1.27778	.47069	.053	-.0105	2.5661
		Between 11 - 20 Years	.30193	.36283	.957	-.6911	1.2950
		More than 20 Years	-.41111	.58284	.981	-2.0064	1.1841
	Between 11 - 20 Years	Less than 3 years	.97585	.45331	.198	-.2649	2.2166
		Between 4 - 10 Years	-.30193	.36283	.957	-1.2950	.6911
		More than 20 Years	-.71304	.56890	.767	-2.2702	.8441
	More than 20 Years	Less than 3 years	1.68889	.64308	.066	-.0712	3.4490
		Between 4 - 10 Years	.41111	.58284	.981	-1.1841	2.0064
		Between 11 - 20 Years	.71304	.56890	.767	-.8441	2.2702

Table: H.11. *ANOVA Test for Teaching Experiences and Usage of Assessment Information*

		Sum of Squares	df	Mean Square	F	Sig.
Make Decision about Individual Student	Between Groups	2.241	3	.747	.602	.617
	Within Groups	65.794	53	1.241		
	Total	68.035	56			
Promotion and Graduation	Between Groups	2.000	3	.667	.797	.501
	Within Groups	44.316	53	.836		
	Total	46.316	56			
Plan Teaching	Between Groups	1.691	3	.564	1.217	.313
	Within Groups	24.554	53	.463		
	Total	26.246	56			
Help Students Address Problems in Learning	Between Groups	1.803	3	.601	1.168	.331
	Within Groups	26.751	52	.514		
	Total	28.554	55			
Assign Grade	Between Groups	.217	3	.072	.090	.965
	Within Groups	41.765	52	.803		
	Total	41.982	55			
Provide Descriptive Feedback	Between Groups	2.449	3	.816	.911	.442
	Within Groups	47.481	53	.896		
	Total	49.930	56			
Rank Student	Between Groups	2.314	3	.771	.636	.595
	Within Groups	64.247	53	1.212		
	Total	66.561	56			
Assessment without Providing Grade	Between Groups	18.004	3	6.001	2.442	.075
	Within Groups	125.341	51	2.458		
	Total	143.345	54			

Table: H.12. *ANOVA Test for Teaching Experiences and Assessment Conceptions*

		Sum of Squares	df	Mean Square	F	Sig.
Improvement	Between Groups	3.077	3	1.026	1.819	.155
	Within Groups	29.882	53	.564		
	Total	32.959	56			
Student_Accountability	Between Groups	2.866	3	.955	1.142	.341
	Within Groups	44.339	53	.837		

	Total	47.205	56			
School_Accountability	Between Groups	2.926	3	.975	.856	.470
	Within Groups	60.423	53	1.140		
	Total	63.349	56			
Irrelevancy	Between Groups	.930	3	.310	.574	.634
	Within Groups	28.605	53	.540		
	Total	29.535	56			

Table: H.13. *Independent Samples Test for Assessment / Measurement Course and Assessment Practices*

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Self-Assessment	Equal variances assumed	1.059	.308	1.618	56	.111	.69915	.43219	-.16663	1.56492
	Equal variances not assumed			1.477	17.323	.158	.69915	.47337	-.29816	1.69645
Individual Project	Equal variances assumed	2.140	.149	.469	56	.641	.11453	.24443	-.37513	.60419
	Equal variances not assumed			.539	24.696	.595	.11453	.21262	-.32364	.55270
Major Exams	Equal variances assumed	.001	.973	-.455	56	.651	-.14701	.32288	-.79382	.49981
	Equal variances not assumed			-.459	19.744	.651	-.14701	.32003	-.81514	.52112
Observation	Equal variances assumed	.556	.459	2.130	55	.038	.88112	.41361	.05223	1.71001
	Equal variances not assumed			1.910	17.054	.073	.88112	.46128	-.09187	1.85411
Group Project	Equal variances assumed	.002	.961	.720	56	.475	.17265	.23978	-.30769	.65299
	Equal variances not assumed			.710	19.115	.486	.17265	.24311	-.33597	.68127
Self-Made Test with Select Type Items	Equal variances assumed	2.379	.129	.346	56	.731	.14872	.42967	-.71202	1.00945
	Equal variances not assumed			.299	16.288	.769	.14872	.49799	-.90545	1.20289
Oral Presentations	Equal variances assumed	.030	.863	.490	56	.626	.14701	.30015	-.45426	.74828
	Equal variances not assumed			.500	20.072	.623	.14701	.29423	-.46660	.76061

Provided by Publishers	Equal variances assumed	4.763	.033	-.816	56	.418	-.27521	.33719	-.95068	.40026
	Equal variances not assumed			-.656	15.236	.522	-.27521	.41949	-1.16814	.61771
Portfolio	Equal variances assumed	.225	.637	.479	54	.634	.18939	.39551	-.60355	.98234
	Equal variances not assumed			.497	18.401	.625	.18939	.38138	-.61062	.98940
Peer Assessment	Equal variances assumed	.821	.369	.220	56	.826	.08718	.39582	-.70574	.88010
	Equal variances not assumed			.235	21.589	.816	.08718	.37062	-.68229	.85665
Questioning	Equal variances assumed	.007	.935	.956	56	.343	.35214	.36845	-.38597	1.09024
	Equal variances not assumed			.877	17.417	.393	.35214	.40171	-.49386	1.19814
Authentic Assessment	Equal variances assumed	.353	.555	1.771	56	.082	.62735	.35423	-.08226	1.33696
	Equal variances not assumed			1.623	17.403	.123	.62735	.38647	-.18659	1.44129
Self-made Test with Supply-type Items	Equal variances assumed	.048	.827	-.757	56	.452	-.23761	.31390	-.86643	.39121
	Equal variances not assumed			-.777	20.261	.446	-.23761	.30581	-.87499	.39978
Essay	Equal variances assumed	6.872	.011	1.481	54	.145	.65909	.44514	-.23336	1.55154
	Equal variances not assumed			1.176	13.759	.260	.65909	.56042	-.54487	1.86305
Performance Assessment	Equal variances assumed	.014	.907	2.873	53	.006	1.07558	.37438	.32468	1.82649
	Equal variances not assumed			2.753	16.671	.014	1.07558	.39077	.24990	1.90126

Table: H.14. *Independent Samples Test for Assessment / Measurement Course and Usage of Assessment Information*

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Make Decision about	Equal variances assumed	1.904	.173	1.205	56	.233	.41368	.34343	-.27430	1.10165

Individual Student	Equal variances not assumed			.983	15.438	.341	.41368	.42094	-.48132	1.30867
Promotion and Graduation	Equal variances assumed	.210	.648	1.420	56	.161	.40000	.28163	-.16418	.96418
	Equal variances not assumed			1.233	16.386	.235	.40000	.32450	-.28659	1.08659
Plan Teaching	Equal variances assumed	.023	.880	.457	56	.649	.10085	.22066	-.34117	.54288
	Equal variances not assumed			.431	18.009	.672	.10085	.23403	-.39081	.59252
Help Students Address Problems in Learning	Equal variances assumed	.014	.907	-.038	55	.970	-.00874	.23204	-.47377	.45629
	Equal variances not assumed			-.037	19.009	.971	-.00874	.23754	-.50591	.48843
Assign Grade	Equal variances assumed	1.368	.247	-.013	55	.990	-.00350	.27597	-.55655	.54956
	Equal variances not assumed			-.011	17.067	.991	-.00350	.30758	-.65225	.64525
Provide Descriptive Feedback	Equal variances assumed	2.618	.111	-.184	56	.855	-.05470	.29723	-.65013	.54073
	Equal variances not assumed			-.137	14.295	.893	-.05470	.40071	-.91248	.80307
Rank Student	Equal variances assumed	11.025	.002	.615	56	.541	.21026	.34215	-.47516	.89567
	Equal variances not assumed			.441	13.959	.666	.21026	.47730	-.81373	1.23424
Assessment without Providing Grade	Equal variances assumed	.502	.482	1.928	54	.059	.96422	.50016	-.03854	1.96698
	Equal variances not assumed			2.049	21.896	.053	.96422	.47067	-.01215	1.94060

Table: H.15. *Independent Samples Test for Assessment / Measurement Course and Assessment Conceptions*

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Improvement	Equal variances assumed	.591	.445	-1.359	56	.179	-.32433	.23858	-.80226	.15360
	Equal variances not assumed			-1.517	23.318	.143	-.32433	.21384	-.76636	.11770

Student_Accountability	Equal variances assumed	.648	.424	-.815	56	.418	-.23533	.28865	-.81356	.34290
	Equal variances not assumed			-.767	17.961	.453	-.23533	.30680	-.87999	.40933
School_Accountability	Equal variances assumed	.054	.817	-2.748	56	.008	-.86467	.31467	-1.49503	-.23432
	Equal variances not assumed			-2.650	18.539	.016	-.86467	.32631	-1.54880	-.18055
Irrelevancy	Equal variances assumed	.136	.714	.791	56	.432	.18051	.22807	-.27637	.63739
	Equal variances not assumed			.785	19.279	.442	.18051	.22983	-.30006	.66109