TITLE: ASSESSMENT OF STUDENTS' LEVELS OF SCIENCE PROCESS SKILLS AND THE EFFECTIVENESS OF INQUIRY-BASED APPROACH AS THE MAIN FEATURES OF THE CURRENT COMPETENCE-BASED CURRICULUM IN TANZANIA.

DISSERTATION
SUBMITTED TO THE FACULTY OF SCIENCE AND MATHEMATICS AT
THE BERGISCHE UNIVERSITY OF WUPPERTAL - GERMANY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF SCIENCE PEDAGOGY

Department of Biology and Didactics

ATHUMAN, JAMAL JUMANNE

FEBRUARY, 2017

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20170405-110001-6
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade3Ahbz3A468-20170405-110001-6-]

Supervisor: Prof'in Dr. Angelika Preisfeld

Referees

1. **Prof’in Dr. Angelika Preisfeld, Bergische University Wuppertal**

2. **Referee: Prof. Dr. Paul W. Dierkes - Goethe-University, Frankfurt**

I. DECLARATION AND COPYRIGHT

I**, Athuman, Jamal Jumanne**, declare that this dissertation is my own original work and that it has not been presented and will not be presented to any other University for a similar or any other degree award.

Signature ………………………

Date……………………………..

II. ACKNOWLEDGEMENT

First of all, I would like to sincerely thank and acknowledge the invaluable contributions, comments and suggestions of my esteemed supervisor Prof'in Dr. Gela Preisfeld of the Faculty of Mathematics Natural and Sciences at the University of Wuppertal. Her encouragement, advice, guidance and quick response made this work possible. I really owe the completion of this work to her. I am very grateful for her patience, motivation, enthusiasm, and immense knowledge in Biology and didactics that, taken together, make her a great mentor. She accepted me as her Ph.D. student without any hesitation when my research proposal was presented to her.

Being a student from a developing country like Tanzania searching for a German Professor for supervision is not a simple exercise. My very special thanks are given to Prof. Dr. Harm Paschen from the University of Bielefeld. He is the one who helped me to find a supervisor when I expressed the desire to pursue my Ph.D. research in one of Germany Universities. His encouragement and help made me feel confident to fulfill my desire and to overcome every difficulty I encountered. At the early stage of my proposal, he helped me to correct grammar mistakes and suggested possible improvements. It is not sufficient to express my gratitude with only a few words to him.

My appreciation should also go to all the organizations which gave me their indispensable generous sponsoring, the German Academic Exchange Service (DAAD) and the Ministry of Education and Vocational Training of Tanzania (MOEVT). Without their financial help, it would not have been possible for me to pursue and to complete this Ph.D. project successfully. I am also extending my heartfelt gratitude to the District Executive Director (DED) of Morogoro municipality for offering me a research permit for data collection. Very special appreciation should go to all teachers and students who participated in this study for their cooperation and inputs without which I could not have obtained my research data. Special thanks should go to the Deputy Vice Chancellor (Academics), on behalf of the Sokoine University of Agriculture (SUA) for offering me a study leave to study in Germany.

A very special appreciation should also go to Dr. Damerau for his invaluable advice in the area of statistics and data analysis. He was always being very supportive of my work. I must also offer specials thanks to my coworkers; Magret Busel, Dr. Monika Steihof, Dr. Sabrina Bleisdell, Dr. Cora Berger, Dr. Sebastian Scherdin, Melanie Beudels, Vadim Kloss, Dr. Stefan Pärschke, Nadine Franken, Nadja Dabbagh, Dr. Viktoria Wittich and Dr. Annika Rodenhauser for their companionship. They have been very wonderful to me. I am also indebted to Habtizel family (Beate, Hannah, Leah and Luka) for the accommodation service they provided me during my three years stay in Wuppertal. Despite our differences in culture, we were able to live very peaceful.

This acknowledgement would be incomplete without mentioning my beloved wife, Saumu A. Mushi who endured my absence for all the time of my study. I appreciate her love, support, inspiration, sacrifice and understanding. My appreciation also goes to my parents Mr. and Mrs. Jumanne Athuman for their moral support. May grace and mercy be bestowed to all of them. I am very much indebted also to my daughters Nuria and Aaliyah and my sons Haarith and Ibrahim who supported me in every possible way to see the completion of this work.

Above all, I owe it all to the Almighty God for granting me the wisdom, health and strength to undertake this research task and enabling me to its completion. Without his grace and love, this work would not have been possible. Lastly, any errors and omissions in this document are solely mine and should not therefore be related to any other person.

III. DEDICATION

This work is dedicated to my lovely wife. Being both a father and mother while I was away was not an easy thing for her. She took every responsibility and suffered all the bitterness to take care of my kids and my family. The work is also dedicated to my daughters Nuriyah and Aaliyah and my sons Haarith and Ibrahim for allowing me to pursue this PhD programme at the expense of their well being.

IV. ABSTRACT

This study was conducted in order to establish a base level of information on whether or not Tanzania students are acquiring competence in science process skills as prescribed in the competence based curriculum of 2005. The competence based curriculum of 2005 was designed to reduce teacher-centered instruction in favor of student-centered learning characterized by active learning, solving problems, challenging existing knowledge, and participating in lively discussion, which is thought to be achieved by an inquiry-based approach . Firstly, the study developed and validated a science process skills test specific for Biology (BPST) to be used in assessing students' competence in this area. In the second stage, the study employed the test that has been developed and validated in the first stage (BPST), to examine the knowledge level of science process skills of advanced level secondary school Biology students in the municipality of Morogoro. Science process skills are one among many competences strongly advocated by the competence based curriculum to learners. In the third stage, the effectiveness of inquiry-based approach on students' scientific process skills development, conceptual understanding of contents and motivation was investigated. Inquiry-based approach to science has also been heavy emphasized by the new curriculum in Tanzania. Eight (08) weeks genetics lessons were designed for a quasi-experimental intervention from Tanzania Biology syllabus on the basis of both inquiry-based learning principles and conventional style.

Through careful attention to the standards for developing validity arguments of a psychometric test, the study provided comparative validity evidence related to test content, response process, and internal structure. Findings from an analysis of data gathered in the pilot study using the developed test (BPST) involving 610 Morogoro students indicated that the test is reliable and valid enough to be employed in a large scale study. The developed Biology process skills test (BPST) had an internal consistency reliability of 0.80 cronbach alpha, a difficulty index of 0.447 and an overall discrimination index of 0.48. Furthermore the content validity of BPST is 0.88, concurrent validity of 0.51 and a construct validity (discriminant correlation coefficient) of 0.34. The readability level of BPST is 72 (fairly easy). The test may also be a useful means of classroom-based research,

evaluation of instruction and learning, curriculum validation, as well as an alternative to authentic methods of assessing scientific skills acquisition.

In the second stage of the study which examined the knowledge level of Morogoro students in the area of science process skills, the validated BPST found that students had a barely average knowledge level of the skills. The mean of test scores was 17.2 out of 35 items in the test which is equivalent to 49.1%. Specifically, Morogoro students performed relative better on items measuring their ability in identifying and controlling variables with score mean of 4.05 (57.8%) out of 07 items and they performed extremely poor on items which measured their skills in analyzing and interpreting data with the mean of 2.34 (33.4%) out of 07 items. Due to the influence of social forces, culture and gender roles in the Tanzania, anecdotal evidence would suggest male students to have higher levels of achievement in science-related disciplines than females. However, the findings from Morogoro Biology students in this study did not support that assertion. Based on the science process skills test scores of the 246 females and 107 males in the study, independent samples t-test found a statistical significant differences in favour of female students.

In the third phase, an analysis of BPST posttest scores revealed that the experimental group students performed better in science process skills after undergoing treatments of inquiry constructivist activities as compared to their counterparts in the control group. An analysis of independent samples t-test based on type of instruction students received at (α) =0.05 produced a p of 0.047 and a t value of 0.633, hence rejecting the null hypothesis (Ho1). However repeated measures ANOVA found that regardless of the method of teaching, there were significant within-groups effects with regard to the development of science process skills. The same result were also obtained with respect to achievement in genetics and motivation. On the otherhand, students scientific skills did not correlate with variables of motivation and genetics. However, generalization of these findings is not possible because of the nature of the study and the sample size used in each stage. It is therefore suggested that replication of the similar study in alternative educational settings is needed before generalization.

TABLE OF CONTENTS

V. LIST OF TABLES

VI. LIST OF FIGURES

VII. LIST OF APPENDICES

VIII. LIST OF ABBREVIATONS AND ACRONYMS

AAAS	American Association for the Advancement of Science
ANOVA	Analysis of Variance
ARCS	Attention, Relevance,Confidance and Satisfaction
ASL	Average Sentence Length
ASW	Average Number of Syllables Per Word
BPST	Biology Process Skills Test
CBC	Competence Based Curriculum
CRT	Criterion-Referenced Testing
CTT	Classical Test Theory
DED	District Executive Director
DNA	Deoxyribonucleic Acid
ENSI	Evolution and the Nature of Science Institutes
FKRA	Flesch-Kincaid Reading Age
FKRS	Flesch-Kincaid Readability Score
FSWEx	Fragebogen zur subjektiv wahrgenommenen Experimentierkompetenz
IBT/IBA	Inquiry- Based Teaching/ Approach
MoEVT	Ministry of Education and Vocational Training of Tanzania
NECTA	National Examination Council of Tanzania
NRT	Norm-Referenced Testing
IOX	Insructional Objective Exchange
RNA	Ribonucleic Acid
SAPA	Science A Process Approach

SEM	Standard Error of Measurement
SMQ-II	Science Motivation Questionnaire II
SPSS	Statistical Package for the Social Sciences
SSP	School Science Project
TIE	Tanzania Institute of Education
TIPS I	Test of Integrated Science Process Skills I
TIPS II	Test of Integrated Science Process Skills II
TSPS	Test of Science Process Skills
TM	Traditional Method of Teaching
URT	United Republic of Tanzania

CHAPTER ONE
GENERAL INTRODUCTION

1.1 Background and context of the study

Across the world, science is increasingly being viewed as a subject of life-long utility to all students, whether or not they enter science-related careers. A more science literate populace is perceived as being better equipped to contribute to the sustainable economic development and to the social welfare (Ware, 1992). Science curriculum innovations and reforms of 1960s to 70s were characterized by attempts to incorporate more inquiry oriented and investigative activities into science classes (Mungandi, 2005; Dillashaw and Okey, 1980). These reforms pushed science programs to start emphasizing the acquisition of science process skills as one of the major goals of science instruction (Padilla, 1990). The aim was to expose students to the world of scientific procedures especially the world of research, experiments, and investigations so that as future scientists, they acquire scientific skills and literacy (Padilla, 1990).

Tanzania also began a process of curriculum reform in the early 2000s, with the goal of transforming Tanzania schooling from exam-oriented education to student-centered learning. Traditional education practices had expected students to passively accept and memorize material presented by teachers, and to reproduce the knowledge on often high-stakes examinations. As a result of these transformations, in 2005 Tanzania came up with the so called 'Competence Based Curriculum' which emphasized among other things, student's competence in science process skills. The new syllabus adopts a two-fold approach of developing students' process skills while testing their content knowledge (URT, 2005). The statements such as students shold be able to compare, classify, use apparatus and equipment, communicate, infer, formulate hypotheses, make prediction, analyse data, define variables operationally are very much seen in the new curriculum (UTR, 2005). These skills are known as scientific process skills and are essential tools for students to explore and acquire scientific knowledge within and outside the classroom (Chiapetta and akaoballa, 2002).

At different times, terms such as scientific method, scientific thinking, scientific inquiry, scientific reasoning and critical thinking have been terms that were used at various times to describe the term science process skills (Mungandi (2005). However, the use of the term science process skills in place of other terms was popularized by the American curriculum project known as Science-A Process Approach (SAPA) which sought to change the emphasis of school science from a mastery of a body of knowledge to the way science was done by scientists (Athuman, 2010; Padilla, 1990 & Mungandi, 2005).

By definition, science process skills are procedural activities that scientists execute when they study or investigate a problem, an issue or a question. Chiappetta and Koballa (2002) define science process skills as a set of broadly transferable abilities appropriate to many science disciplines and reflective of the behavior of scientists. Bilgin (2006) on the other hand defined science process skills as an understanding of methods and procedures of scientific investigation. They are hierarchically organized, ranging from the simplest to the more complex higher order ones, called integrated science process skills (Dyer et al. 2004). Integrated science process skills include skills in formulating hypotheses, identifying and controlling variables, defining operationally, experimenting, and interpreting data (Chiappetta & Koballa, 2002). Basic science process skills, on the other hand, are designed to provide a foundation for the learning of integrated process skills (Chiappetta & Koballa, 2002). They include skills in observing, measuring, using numbers, classifying, seriating, predicting, and inferring (Hamilton & Swortzel, 2007).

Science process skills are necessary for dealing with everyday life and in developing an understanding of the natural world. According to Harlen (2000) these skills are the necessary means by which learners engages with the world and gains intellectual control of it through the formation of concepts and development of scientific thinking (Harlen, 2000). Science process skills contributes to students´ scientific literacy with their emphasis on hypothesizing, manipulating the physical world and reasoning from data(Chiappetta & Koballa, 2002). According to Harlen (1999), science process skills allow students to tie new information to old information. Students gradually build small facts together to produce a larger understanding of the concept, critical thinking, and scientific reasoning skills (Pratt & Hackett, 1989).

Apart from emphasizing on the need for the acqusition of science process skills, the shift from content to competence-based education in Tanzania also involved some pedagogical changes. This involves incorporating outcome-based learning and constructivist philosophies rather than a theoretical understanding of concepts as it was in the traditional content-based curriculum (Tilya & Mafumiko, 2008). This is a shift away from the learning of more-or-less isolated facts and facets in biology, chemistry, and physics towards a restructuring of science teaching along the general principles of the respective science domains (Tilya & Mafumiko, 2008). The changes included also the addition of dimensions such as problem-based learning, understanding the basics of the nature of science, and engaging students in the methods of science (URT, 2008). Since then, practical work and science process skills, in general, has been solidly built into Tanzania science syllabuses (Osaki, 2007).

Inquiry based instruction is a teaching strategy that aims to develop students' skills to deal with problems that they may encounter by using the methods used by scientists via researching, investigating, analyzing and inquiring in the classroom (Crawford, 2000). The new curriculum require Tanzania students to learn scientific subjects such as Biology, Physics and Chemistry in the same way science is done scientists through inquiry constructivist approach. Results of several studies (eg Crawford, 2000; Carin & Bass, 2001; Kyle, Bonnstetter, & Gadsden, 1988; Brew, 2003) have shown that students' scientific process skills can be developed by using inquiry or investigative approach of teaching and learning science that gives them opportunities to practice these skills. According to the new curriculum, science education for the future involves teaching students more than just the basic concepts of science. Students need to be equipped with the skills to be able to use scientific knowledge to identify questions, and to draw evidence-based conclusions in order to understand and make decisions about the natural world and the changes made to it through human activity (URT, 2005).

1.2 The 2005 competence based curriculum (CBC) features related to this study

In the early 2000s, Tanzania began a process of curriculum reform with the goal of transforming Tanzania schooling from exam-oriented education to student-centered learning. Traditional education practices had expected students to passively accept and memorize material presented by teachers, and to reproduce the knowledge on often high-stakes examinations. As a result of these transformations, in 2005 Tanzania came up with the so called 'Competence Based Curriculum' which emphasized among other things, students' competence in science process skills. The curriculum emphasized the need of Tanzania science students to learn science subjects such as Biology, Physics and Chemistry in the same way science is done scientists.

By 2005 the government through the Tanzania Institute of Education (TIE) had already completed the process of revising the curriculum of primary, secondary, and teacher education levels from that of content-based to competence-based paradigm (URT, 2005). The new 2005 Competence Based Curriculum was streamlined to address the needs of developing analytical and market-oriented skills to learners (URT, 2008). According to the United Republic of Tanzania (2006), the Ministry of Education planned not only to review the existed curriculum but also to orient teachers on the requirements of the new curriculum and strengthen the provision of teaching and learning materials. The new curriculum emphasizes the need for Tanzania science students to acquire science process skills and the need of science teachers to use inquiry-based approach in science teaching. These two features in the new curriculum are the center of interest in this study as discussed in the section 1.2.1 and 1.2.2.

1.2.1 Competence in science process skills

The Tanzania Institute of Education (TIE) identified science process skills as being essential in creating the 2005 competence based curriculum (MoEVT, 2005). The curriculum has incorporated these skills both in scientific investigations and in construction science knowledge of science curriculum. As a result of this move, many of the science syllabuses, guides, reference books and instructional materials for the revised curriculum acknowledge the need for science process skills acquisition. The revised secondary school science syllabuses explicitly state and emphasize the need for science learners to acquire competence in science process skills. The new ordinary level secondary school Biology syllabus of 2005 for example, has the following competence objective statements;

i. Students should have the ability to plan, record, analyze and interpret data from scientific investigations using appropriate methods and technology to generate relevant information in biological science.
ii. Students should be able to develop necessary biological practical skills.
iii. Students should have the ability to apply scientific skills and procedures in interpreting various biological data (p.ii-v).

In addition, the syllabus (p.1) stipulates that science process skills should start as early as from form one when a learner has just started secondary education. The Biology syllabus for example states that, at the end of the year, a form one student should be able to; (i) develop and apply basic knowledge and skills on scientific processes of studying Biology and (ii) develop mastery of carrying out experiments on various biological processes (p.1).

Science process skills also reappear in the list of objectives of higher classes and in the list of other science subject syllabuses. For example, a new secondary Chemistry syllabus of 2005 maintains that students should be able to, (i) think critically and evaluate scientific procedures (ii) synthesize, analyze, and communicate scientifically (iii) design and carry out experiments to prove a mastery of scientific procedures, etc (URT, 2005). All these learning abilities and competences to be acquired by learners are collectively known as science process skills (Chiappetta & Koballa, 2002).

1.2.2 The competence based curriculum and the use of inquiry-based approach to science

Current perspectives on science teaching and learning, as well as the current 2005 curriculum policy in Tanzania, stresses on the use of teaching methods that promote active engagement of learners during teaching and learning processes. This curriculum was reviewed in the spirit of constructivism to enhance participatory and inquiry approaches to teaching (Tilya & Mafumiko, 2008). With constructivism philosophy, learners are encouraged to participate actively in the lesson, use their pre-concept knowledge, and engage in classroom activities so as to construct meaning out of the lesson (Kelly, 1991). The new curriculum policy acknowledges the fact that, inquiry-based teaching approach must be an integral part of science education if science process skills are to be acquired by students. In the advanced level Biology syllabus of Tanzania of 2010 for example it is stated that…

……. Teachers are advised to use participatory teaching and learning strategies as much as possible to help learners demonstrate self-esteem confidence and assertiveness (Pg.vii).

As one of the constructivist participatory methods of teaching, the inquiry-based approach requires teachers to facilitate the inquiry process, granting student responsibilities for their learning while modeling and scaffolding the cognitive and investigative processes involved (Lebow, 1993; Myer, 2004; Kirschner et al. 2006). The approach provides opportunities to understand the scientific inquiry process and to develop general investigative abilities (such as posing and pursuing open-ended questions, synthesizing information, planning and conducting experiments and analyzing and presenting results), as well as to gain deeper and broader science content knowledge that has real-world application (Prawat, & Floden, 1994). The skills are collectively called Science Process skills. In the teaching of science through inquiry approach, teachers act as facilitators, motivators, and inspires for students in driving the lesson. This is in contrast to a traditional paradigm where teacher´s role is to decide, control and direct student learning in what is known as banking education (Barakatas, 2005). The teacher is an authority who decides what and how their students should be teaching (Chung, 2004). Lessons are designed with a view to specific learning outcomes which are outlined in structured lesson plans. Evaluation of learning is based on student performance on objective tests (Floresc & Kaylor, 2007).

Therefore in addition to the construction of a valid science process skills test for assessment of Tanzania Biology students' competence in the area of science process skills, the study investigated the effectiveness of inquiry-based teaching approach on the development of students science process skills, conceptual understanding and motivation of students. A review of the literature by the researcher failed to found a single study in Tanzania on inquiry-based teaching with a special focus on its effectiveness in promoting conceptual understanding of contents, science process skills development and in enhancing motivation.

1.3 The need for constructing and validating a test for science process skills in the context of Tanzania education system

Having established the importance of science process skills in students scientific thinking and knowledge acquisition (section 1.1) and the extent to which the revised Tanzania curriculum have addressed them (section 1.2), the question that arises is, to what extent have students who use this curriculum and the related instructional materials such as syllabuses, reference books, and guides acquire science process skills as prescribed?. The answer to this kind of a question according to Mungandi (2005) lies in the effective assessment of science learners' competence in these skills. A review of literature and studies in Tanzania shows not much work if any have been done in this area of test construction and validation for assessing scientific skills. Some researchers (Wiggins, 1989; Gronlund, 1998; Hofstein & Lunetta, 2004) have proposed an authentical based assessment of science process skills. Wiggins (1989) defined authentic assessment as the kind of assessment which involves real-life tasks, performances, or challenges that replicate the problems faced by a scientist, or expert in a particular field. It involves real tasks rather than drills, worksheets, or isolated multiple choice questions. In the context of students´performance in science labs, Gronlund (1998) argue that if you want to determine whether students can conduct an experiment, let them conduct an experiment.

However, authentic methods of assessing science process skills competence such as through laboratory practical work have a number of constraints particularly in the context of teaching large under-resourced science classes (Onwu & Stoffels, 2005; Mungandi, 2005). This is particularly true of Tanzania secondary schools. A survey conducted by Tanzania researchers (Osaki, 2007; Osaki & Njabili, 2004; Semali & Mehta, 2012; Athuman, 2010) on the teaching of science in selected schools in Tanzania reveals science classes being characterized by a large number of students. Overcrowding of students made effective guiding of science practicals and authentic assessment impossible. Moreover, most of these schools, especially community owned secondary schools either do not have science laboratories or they are poorly equipped with of reagents, apparatus, and samples (Athuman, 2010; Osaki & Njabili, 2004). The fact above compound to the problem of relying on and make authentic assessment almost impossible in many schools. Objective tests in the multiple choices format is an alternative in measuring students science process skills.

In the situation where authentic assessment of scientific skills is impossible, the use of multiple choice questions became an alternative. This is the common way that has been used worldwide to assess science process skills, especially in large under-resourced science classes. Multiple choice tests through the use of paper and pencil test do not require laboratories and expensive resources. Review of science education literature failed to identify any study in Tanzania that come up with a with a science process skills test for Tanzania students. Although a number of science process skills test exist worldwide, they are not suitable for Tanzania learners. This is because they have been developed and validated outside Tanzania hence not taking into consideration internal realities within education system settings. Another problem with these international science process skills tests is the fact that they are not knowledge domain specific. Some researchers tend to believe that scientific inquiry is independent of domain knowledge (Nehring et al. 2012; Millar, 1987; Lock, 1993). The current study was based on the assumption that students' ability to use process skills partly depend on the extent of their content knowledge they are asked to work on. Hence, Tanzania science teachers need a convenience, objective, and a cost-effective means of assessing science process skills to supplement the use of the science laboratory practical method. Development and validation of a science process skills test was one the aims of this study.

1.4 The rationale of genetics literacy to Tanzania students and its inclusion in the competence based curriculum of 2005

Conceptual understanding of genetics is one of the key issues addressed by the competence-based curriculum of 2005 in Tanzania. The topic forms one of the central core contents of advanced level Biology contents. Genetics is defined by Jennings (2004) as a field of study that is concerned with heredity and how particular qualities or traits are passed on from parents to offspring. The term genetics literacy was proposed as a part of scientific literacy to emphasize the issues and challenges that are related to genetics and biotechnology (Jennings, 2004; Freidenreich et al. 2011). Genetics literacy provides sufficient knowledge and appreciation of genetics principles to allow informed decision-making and for personal well-being and effective participation in social decisions on genetics issues (Bowling et al. 2008).

Over the last several decades, the role of genetic technologies in health and public policy has persistently increased (Miller, 1998) and new knowledge in genetics continues to have significant implications for individuals and society (Tsui and Treagus, 2010; Lewis & Kattman, 2004). Rapid advancements in genetics and genetic technology are creating opportunities for the understanding, prevention, treatment and cure of human diseases. Tsui and Treagust (2010) stressed the importance of having contemporary knowledge on DNA, genes, and their relations to human affairs on making informed decisions about ethically and socially controversial issues. Genetic issues now play a large role in health and public policy (Miller 1998 & Freidenreich et al. 2011). Competence in genetics is necessary not only to make thoroughly informed decisions about socio-scientific issues such as cloning, genetic screening, gene therapy and genetically modified foods but also their ethical, legal, and social implications (Bowling, 2007). Poor genetic literacy for example in Tanzania has led to the brutal murder and attacks on innocent men, women, and especially children with albinism under the influence of witchcraft and superstition and desperation for wealth. These misconceptions, coupled with the lack of education are some of the key reasons that albinism is so heavily persecuted. Enhancing students´ understanding of genetics can improve communication regarding genetic information and technologies, and help to ensure its appropriate use (Tsui & Treagus, 2010; Lewis & Kattman, 2004).

Genetics has been chosen as a focus point in this study because it is a topic that offers a lot of opportunities where students can practice realistic problem solving. Genetics is one of those topics that are relevant to our daily lives. Understanding how genetics plays a role in our past, present and future helps us to better understand ourselves and those around us. Available studies reported that genetics is among the main topics that students struggle with serious conceptual difficulties (Duncan & Reiser, 2007; Jennings, 2004; Lewis & Kattman, 2004) therefore genetics topic is crucial to be selected as a case study in assessing which one is the effective instructional method between conventional methods or an inquiry-based approach to science. Although the problem-solving skills gained in genetics relate to a specific domain of learning, one hopes the skills gained in learning how to approach problem-solving in genetics would be transferable to other areas of life. According to the advanced level Biology syllabus of Tanzania (2010), genetic contents are categorized into the following subtopics i. Hereditary materials (DNA/RNA), ii. Genetic coding and protein synthesis, iii. Mendelian inheritance and pedigree, iv. NonMendelian inheritance v. Sex linked inheritance and vi. Gene and chromosomal mutation.

1.5 Problem statement

It is eleven years now since the inception of the competence-based curriculum in Tanzania. The newly revised competence based curriculum of 2005 has placed a heavy emphasis on the need for secondary school science learners to acquire science process skills. The curriculum emphasized the need of Tanzania science Students to learn scientific subjects such as Biology, Physics and Chemistry in the same way science is done scientists. Despite such a dramatic shift in curriculum policy, little is known about whether or not the reform efforts are truly transforming the educational experiences of students. There is no clear evidence of whether or not learners are appropriately acquiring competence in these scientific skills as prescribed in the curriculum. According to Berliner (1986), successful implementation of a curriculum reform should be measured by the extent to which learners have acquired the targeted objectives. The learner is the primary reason for developing or reforming any curriculum. Plowden (1967) also is convinced that…………..

> *"At the heart of the educational process lies the child hence the evaluation of a curriculum reform must begin with learners" (Plowden 1967:7).*

A review of the literature failed to identify any research that has investigated whether or not Tanzania students are acquiring competence in science process skills as planned in the curriculum. A review of literature also failed to identify any study that has attempted to construct and validate a scientific test for measuring science process skills in the context of Tanzania. On the other hand, most of the constructed tests for science process skills are non-discipline specific. They tend are consisted of Biology, Chemistry and Physics items all together (see test of integrated process skills by Dillashaw and Okey (1980), integrated science process skill test (TIPS II) by Burns et al. (1985), process skills test by McKenzie and Padilla (1986), Science process skills test by Onwu and Mozube (1992), integrated process skill by Mungandi, 2005 and the recent one by Shahali & Halim (2010). Therefore with curriculum reforms in Tanzania and the absence of a valid measuring tool, it becomes vital for this study to come up with a valid tool and use the tool to examine students' level of science process skills. Moreover, it is unknown whether there is a statistical difference in term of

Tanzania students’ performance in individual integrated science process skills. Therefore it is necessary for this study to come up with the test that will have an equal number of questions for measuring pupils' scientific skills in i. formulating hypotheses, ii. defining variable operationally, iii. Identifying and controlling variables, iv. planning investigations as well as questions on v. analyzing and interpreting data.

For students’ acquisition of science process skills, the inquiry-based approach must be an integral part of science teaching. The ministry of education and vocational training in Tanzania (2008) recommend that science instruction and learning should be well grounded in inquiry. As seen in section 1.3 above, inquiry-based participatory methods of science teaching have been acknowledged in the competence-based syllabuses such as the syllabus for advanced level Biology

…….. Teachers are advised to use participatory teaching and learning strategies as much as possible to help learners demonstrate self-esteem confidence and assertiveness (Pg.vii).

Despite numerous studies on the value of inquiry teaching approach worldwide and its acknowledgment in the Tanzania syllabuses, review of literature and studies failed to identify any study that scientifically investigated the effectiveness of the approach on students’ scientific process skills development, conceptual understanding of Biology contents and motivation towards science process skills. Hence it became vital again for this study to develop genetics lesson modules based on inquiry teaching and learning principles, implement to students and measure its effectiveness in science process skills development, conceptual understanding of the topic and motivation levels of students as compared to the conventional approaches. Genetics is a focal point because the topic offers a lot of opportunities where students can practice realistic problem solving making it suitable for inquiry-based practices.

1.6 Objectives of the study

1.6.1 General objective of the study

This study was a threefold i. developing and validating a test that will be used in assessing the competence of advanced level Biology students in the area of science process skills ii. employing the validated test to examine the knowledge level of science process skills of advanced secondary school Biology students in Morogoro municipality. iii. Examining the effectiveness of inquiry-based approach on the development of students' science process skills, conceptual understanding and motivation of students towards science process skills

1.6.2 Specific objectives of the study

Specifically, the study intended to;

i. Construct and validate a science process skills test specific to Biology using Tanzania syllabus for measuring the level of skills of advanced level Biology students.

ii. Examine the knowledge level of science process skills of advanced level secondary school Biology students in the selected schools in Morogoro municipality using the constructed and validated test.

iii. Determine whether there a statically significant difference in the performance of the science process skills of Morogoro Biology students based on their sex and grade level.

iv. Compare and describe the performance of Biology students in Morogoro municipality in each of the five integrated science process skills under study (formulating hypotheses, defining operationally, and controlling variables, designing experiments and interpreting data).

v. Assess the effectiveness of inquiry-based approach on the development of students' science process skills in the selected schools in the municipality of Morogoro by comparing it with the conventional approach.

vi. Examine the effectiveness of inquiry-based approach on students' conceptual understanding of Biology contents (genetics being the case study) by comparing it with the conventional direct approach.

vii. Compare the effectiveness of inquiry-based approach on the development of students' motivational constructs (intrinsic motivation, grade motivation, career motivation, self-efficacy, self-determination, and self-concept) towards science process skills.

viii. Determine the correlation existing between the performance of students in the process skills test with their achievement in genetics test and motivation towards science process skills

1.7 Hypotheses and questions guiding the study at different stages

1.7.1 Research questions in the first stage

The first stage of this study aimed at developing and validating the test of integrated science process skills specific to Biology (BPST) was guided by the following questions.

i. What are experts' opinions towards the face validity and content validity of the developed science process skills test?
ii. What is the difficulty index and discrimination index of each item in the developed process skills test?
iii. How effective is each distracter in the multiple choice items of the developed test?
iv. What is the internal consistent reliability and standard error of measurement of the developed science process skills test?
v. What is the calculated construct validity and the concurrent validity of the developed science process skills test?
vi. What is the calculated readability score and readability grade level of the developed test as a text?

1.7.2 Hypotheses and questions in the second stage

The second stage employed the validated test instrument (BPST) to measure Tanzania students' level of science process skills. The stage was guided by the following research questions

i. What is the general performance of advanced level Biology students in the in science process skills test validated in the selected schools in Morogoro municipality?

ii. What is the performance of Biology students in the test items measuring their ability in formulating hypotheses, defining variable operationally, identifying and controlling variables, planning investigations, as well in analyzing and interpreting data?

iii. Is there a difference in the performance of Biology students in Morogoro across individual integrated science process skills under study?

1.7.3 Hypotheses in the third stage (the intervention phase)

The third stage was an intervention phase and it intended to compare the effectiveness of inquiry-based teaching approach and conventional method on students' scientific process skills development, conceptual understanding and motivation towards science process skills. This quasi - experimental phase was guided by the following hypotheses

i. Ho1: There is no statistically significant difference in students' science process skills achievement between those exposed to inquiry-based teaching (IBT) and those exposed to traditional methods (TM).

ii. Ho2: There is no statistically significant difference in the conceptual understanding of genetics between students exposed to inquiry-based teaching (IBT) and those exposed to traditional methods (TM).

iii. Ho3: There is no statistically significant difference in i. intrinsic motivation ii. grade motivation, iii. career motivation, iv. self-efficacy, v. self-determination, and vi. self-concept towards science process skills between students exposed to inquiry-based teaching (IBT) and those exposed to traditional methods (TM).

iv. Ho4: There is no significant correlation between students achievement science process skills and their i. achievement in genetics knowledge ii. intrinsic motivation iii. grade motivation, iv. career motivation, v. self-efficacy, vi. self-determination, and vii. self-concept towards science process skills.

1.8 Research tasks involved in each stage of the study

The study was conducted in three interrelated stages with each stage having unique tasks involved.

1.8.1 Tasks in developig and validating a test of integrated science process skills

Tasks Involved

i. Identifying the objectives of the test, specifying the contents to be involved and forming a specification table delineating indicators and the number of items
ii. Writing appropriate test items that match the advanced level Biology syllabus contents of Tanzania and the table of specification
iii. Checking the test with a team of experts to obtain the evidence of its content validity and agree on its face validity
iv. Conducting a pilot study with the test in the selected secondary schools in Tanzania
v. Using the obtained results to conduct item analysis in order to establish test's internal consistence reliability as well as the difficulty and discrimination indices.
vi. Modifying and rejecting some items that have difficulty and discrimination indices outside the acceptable range of acceptance
vii. Retesting the tool to the same group of students and obtain the final draft

1.8.2 Tasks in measuring students knowledge of science process skills

i. Administering the test to the selected sample of advanced level Biology students in Morogoro municipality
ii. Marking and analyzing science process skills test results conducted
iii. Assessing the performance of advanced level Biology students in each of the five integrated science process skills under study (formulating hypotheses, defining variable operationally, identifying and controlling variables, planning investigations as well as interpreting data)
iv. Determine whether there is a statistical difference in terms of performance in the test across individual integrated science process skills under study).

1.8.3 Tasks in assessing the effectiveness of inquiry-based approach on process skills development, conceptual understanding and motivation

i. Designing genetics modules as per Tanzania Biology syllabus sub-topics on the basis of inquiry-based teaching and learning principles for intervention
ii. Designing and moderating a genetics test for measuring students conceptual understanding as a result of teaching intervention
iii. Adopting and employing Science Motivation Questionnaire II (SMQ-II) developed by Glynn et al. (2011) to pretest students motivation ((intrinsic motivation, grade motivation, career motivation, self-efficacy, and self-determination) towards science process skills, FSWEx questionnaire by Kastern´s (2012) to pretest students self-concept
iv. Employing science process skills test to pre-test students level of science process skills, and genetics test to pre-test students conceptual understanding before teaching intervention
v. Teaching the designed genetics inquiry-based modules to students in the sampled schools in Morogoro municipality
vi. After the end of the inquiry-based intervention, employing science process skills test, cognitive test, and Likert scales to obtain posttest results
vii. Analyzing and examining changes in the performance of students' science process skills, conceptual understanding of genetics, motivation if any.

1.9 THEORETICAL AND CONCEPTUAL FRAMEWORK OF THE STUDY

1.9.1 Theories and approaches that formed the conceptual framework at the first stage of the study involving test construction and validation

The first stage of this study was guided by a number of theories, concepts and approaches related to science process skills, test construction and validation. These theories included i. science process skills approach by Chiappetta and Koballa (2002), ii. the Bloom's taxonomy theory of educational objectives iii. and the criterion-referenced testing (CRT) assessment model. On the other hand, procedures suggested Spector (1992) guided the process of developing and validating the Biology process skills test (BPST). In item selection, the study was guided by Instructional Objective Exchange (IOX) theory and a table of specification concept. Lastly, the classical test theory provided the theoretical framework for analyzing students item responses after a pilot study in an attempt to validate items selected. The following section briefly explains each theory and concept that shaped the framework of thinking during test construction and validation.

1.9.1 Science process approach Chiappetta and Koballa (2002)

The science process skills approach was developed by Chiappetta & Koballa (2002) which focuses on the teaching and learning of science process skills to secondary school students. Chiappetta & Koballa (2002) defines science process skills as a set of broadly transferable abilities that are appropriate to many science disciplines and are the reflective of the behavior of scientists. The approach classifies science process skills into two major groups; Basic and Integrated science process skills. Basic science process skills include skills in observing, inferring, measuring, classifying, and predicting. They provide a foundation for the acquisition of the higher order complex skills called integrated science process skills. The integrated science process skills include skills in identifying and controlling variables, defining variables operationally, formulating hypotheses, interpreting data and in designing experiments (Chiappetta & Koballa, 2002). This study focused on integrated science process skills because the test developed was meant for advanced level secondary school students. Skills and their definitions are presented in Table 1.1 below.

Table 1.1 Classification of science process skills into basic and integrated skills by Chiappetta and Koballa (2002)

Process Skill	Definition
i. Basic Science Process Skills	
Observing	Noting the properties of objects and situations using the five senses
Classifying	Relating objects and events according to their properties or attributes
Space/time relations	Visualizing and manipulating objects and events, dealing with shapes, time, distance and speed
Using numbers	Using quantitative relationships
Measuring	Expressing the amount of an object or substance in quantitative terms
Inferring	Giving an explanation for a particular object or event
Predicting	Forecasting a future occurrence based on past observation or the extension of data
ii. Integrated science process skills	
Defining operationally	Developing statements that present concrete descriptions of an object or event by telling one what to do or observe
Formulating models	Constructing images, objects, or mathematical formulas to explain ideas
Controlling variables	Manipulating and controlling properties that relate to situations or events for the purpose of determining causation
Interpreting data	Arriving at explanations, inferences, or hypotheses from data that have been graphed or placed in a table
Hypothesizing	Stating a tentative generalization of observations or inferences that may be used to explain a relatively larger number of events but that is subject to immediate or eventual testing by one or more experiments
Experimenting	Testing a hypothesis through the manipulation and control of independent variables and noting the effects on a dependent variable; interpreting and presenting results in the form of a report that others can follow to replicate the experiment

From Chiappetta, E. L., & Koballa, T. R., Jr. (2002). *Science instruction in the middle and secondary schools.* Upper Saddle River, N.J: Merrill Prentice Hall.

1.9.2 Bloom's taxonomy of educational objectives theory (1956)

In connection to the science process skills approach, Bloom's taxonomy of learning objectives (1956) provided another dimension of thinking to this study at this stage. Bloom's taxonomy is a set of three hierarchical models (cognitive, affective and psychomotor) used to classify educational learning objectives into levels of complexity and specificity (Huitt, 2011). The taxonomy provides a systematic way of describing how a learner's performance develops from simple to complex levels in their affective, psychomotor and cognitive domains of learning (Krathwohl et al. 2001; Huitt, 2011). In the cognitive domain which is the focus of this study, there are six stages in increasing complexity namely knowledge, comprehension, application, analysis, synthesis and evaluation. Knowledge, comprehension, and application are referred to as lower order levels while analysis, synthesis, and evaluation are considered to be higher order abilities (Huitt, 2011). Lower order abilities are a requisite and foundation for the higher order ones as shown in table 1.2 below.

Table 1.2 Bloom's (1956) taxonomy of the cognitive domain

	Cognitive Domain Levels	Verbs Used for Objectives
Lowest level	Knowledge	define, memorize, repeat, record, list, recall, name, relate, collect, label, specify, cite, enumerate, tell, recount
	Comprehension	restate, summarize, discuss, describe, recognize, explain, express, identify, locate, report, retell, review, translate
	Application	exhibit, solve, interview, simulate, apply, employ, use, demonstrate, dramatize, practice, illustrate, operate, calculate, show, experiment
Higher levels	Analysis	interpret, classify, analyze, arrange, differentiate, group, compare, organize, contrast, examine, scrutinize, survey, categorize, dissect, probe, inventory, investigate, question, discover, text, inquire, distinguish, detect, diagram, inspect
	Synthesis	compose, setup, plan, prepare, propose, imagine, produce, hypothesize, invent, incorporate, develop, generalize, design, originate, formulate, predict, arrange, contrive, assemble, concoct, construct, systematize, create
	Evaluation	judge, assess, decide, measure, appraise, estimate, evaluate, infer, rate, deduce, compare, score, value, predict, revise, choose, conclude, recommend, select, determine, criticize

Source: Huitt, W. (2011). Bloom et al.'s taxonomy of the cognitive domain. *Educational Psychology Interactive*. Valdosta, GA: Valdosta State University. Retrieved from http://www.edpsycinteractive.org/topics/cognition/bloom.html

Demonstration of integrated science process skills is said to require the use of higher order thinking skills since competence in science process skills entails the ability to apply the learnt materials to a new concrete situation, analyzing relationships between parts and recognize of organizational principles involved, then synthesizing parts together to form a new whole and then evaluating or judging the value of materials such as judging the adequacy with which conclusions are supported by data (Mungandi, 2005).

The close relationship between science process skills and higher order thinking skills is acknowledged by several researchers (Mungandi, 2005). For instance Padila et al. (1983) in their study titled 'The relationship between science process skills and formal thinking abilities' found that formal thinking and science process abilities are highly interrelated. Furthermore Baird and Borick (1987) in their validity considerations for the study of formal reasoning and integrated Science process skills concluded that formal reasoning and integrated science process skills competence share more variance than expected and that they may not comprise distinctly different traits. Based on this theory, the items used in developing the Biology process skills test were higher level questions. These questions were consistent with analysis, synthesis and evaluation levels in Bloom´s cognitive domain.

1.9.2 Assessment, construction, validation and item response theories

Apart from the science process skills approach by Chiappetta and Koballa (2002) and Bloom's taxonomical theory on the domains of education objectives (1956), several assessment, test construction, validation and item response analysis models also formed the theoretical and conceptual framework for this study during the development and validation of the science process skills test.

Firstly in assessment, norm-referenced testing (NRT) and criterion-referenced testing (CRT) represent two of the main assessment models, which have historically been used in education (Bond, 1996; Rivera, 2007). A criterion-referenced test is a test that provides a basis for determining a candidate's level of knowledge and skills in relation to a well-defined domain of content. It measures specific skills and knowledge, which make up a designated curriculum

(Huitt, 1996). NRT on the other hand measures the performance of a group of test takers against the performance of another group of test takers (Huitt, 1996). CRT offers many advantages over NRT (Bond, 1996; Rivera, 2007). The main aim of CRT is to determine whether each student has achieved specific skills or concepts identified by teachers and curriculum experts. Each skill is expressed as an instructional objective. Hence CRT provides more meaningful data and a more accurate interpretation of performance. It assesses competency on certification exams, evaluates programs, and monitors an individual's progress or deficiency in objective-based instruction (Huitt, 1996). The science process skills test that was constructed and validated in this study is a typical criterion referenced type as it intends to measure scientific skills as stipulated in the Tanzania curriculum for high school students.

In the area of test construction and development, Popham and Husek (1969) were among the first researchers to write about technical matters associated with building criterion-referenced tests. They also offered a set of methods for interpreting test scores referred to as objective-based or domain-based measurement (Rivera, 2007). Since then, there have been hundreds of research papers related to CRT construction. For example, Hambleton & Simon (1980) offered six stages as (1) writing objectives, (2) preparing and validating test items, (3) determining test lengths (4) selecting test items, (5) assessing the reliability and validity of test scores and decisions, and (6) evaluating tests. Similar to Popham (1975), Roid and Haladyna (1982) described a five-step process as ideal for developing CRT. Recently, Shock and Coscarelli (2000) developed a 13-step systematic model to follow. Burton & Mazerolle (2011) recommended four steps. Step one consists of defining constructs and determining domain content. Step two involves generating items for the survey and judging the appropriateness of the items. Step three is to design and conduct studies to test the scale. Lastly, step four involves finalizing the scale based on data collected in the third step (Burton & Mazerolle, 2011). In this study however, steps as suggested by Spector (1992) were followed throughout the process of developing and validating the Biology process skills test (BPST). They include i. deciding on the principles upon which the constructed test will be based, ii. defining constructs and objectives to be measured, iii. collection and preparation of test items , iv. conducting initial validation, v. carrying out pilot testing, vi. performing item analysis, vi. establishing reliability evidence, vii. calculating the validity of test scores, viii. estimating readability index of the test, and finally x. preparing the final valid test.

On the other hand, a number of procedural methods and methods based on substantive theories have been offered by measurement specialists to be used in developing criterion-referenced items (Alkin, 1969). During the early 1960s, the Instructional Objective Exchange (IOX) theory was a famous theory specialized in using an objectives-based approach to developing criterion-referenced measurements (Alkin, 1969). The proponents of the theory believe that it will be easier for the busy teacher or administrator to select from among objectives, and to generate only a very few, than it would be for him to formulate an entire set of behavioral objectives and measurement items (Alkin, 1969). Using item forms (Hivey et al. 1968), IOX was able to develop amplified objectives. The amplified objectives were not widely utilized so IOX found success by delimiting the amplified behavioral objectives and developing test specifications (Rivera, 2007). A table of specifications is a two-way chart which describes the topics to be covered by a test and the number of items or points which will be associated with each topic. The purpose of this table is to identify the achievement domains being measured and to ensure that a fair and representative sample of questions appears on the test (Chase, 1999). It provides the teacher with evidence that a test has content validity that it covers what should be covered as it is not possible to assess everything a teacher has taught in a single test (Chase, 1999). Hence under the framework of Instructional Objective Exchange (IOX) theory, many items for the science process skills test in this study were collected from different sources (Books, past papers) under the guidance of the specification table. There are a number of concepts, rules, and principles, a technique that can be applied to almost any discipline (Tiemann & Markle, 1983). However, there is no concrete rule book that instructs teachers or researchers on how to generate criterion-referenced measurements, though Hambleton and Rogers (1991) offered the most detailed steps including how to generate good multiple-choice test items. Other models for item generation include the theory of mapping sentences which is based on the structural facet theory (Guttman, 1959), and the factor-based construction method, which generates items through factor-analysis (Meeker & Roid, 1985).

However, writing the test item does not produce an item ready to be tested until it is validated. In item validation, the study was guided by Hambleton (1994) who provided features to focus on when reviewing a CRT item's content. Hambleton (1994) offered guidelines and methods used to review and validate items. On the other hand, there have been multiple techniques established for reviewing item-objective congruence based on large-scale assessments such as expert judgment used to calculate the index of item-objective congruence, a rating of the item-objective match on a scale conducted by experts, and the use of a matching task (Hambleton, 1994). Hence in this study, a rating of the item-objective match on a scale conducted by experts was adopted through creating a content validation form in order to determine content validity of each item and the face validity of the test.

The classical test theory provided the theoretical framework for analyzing students item responses after the pilot study. Classical test theory (true score theory) is a psychometric theory which assumes that each person has a true score, T, that would be obtained if there were no errors in measurement (Traub, 1997). A person's true score is defined as the expected number-correct score over an infinite number of independent administrations of the test. Lord and Novick (1968) introduced classical test theory (CTT) approaches to the behavioral sciences. They introduced the classical linear model and its application to estimating parameters such as true score and error variances of latent trait variables (Rivera, 2007). Common statistics are used to describe CTT parameters including p-values, item discrimination, point-biserial correlation coefficient, alpha coefficient, and variance (Rivera, 2007). Analysis of these parameters provides evidence for the validity of criterion-referenced examinations. CTT statistics can also be used to determine values of reliability through the use of internal consistency methods such as split halves (Traub, 1997) and item co-variance. Other methods used in determining internal consistency reliability apart from split half techniques are Kuder-Richardson 20 (KR 20) and Kuder-Richardson 21 (KR 21) along with Cronbach's alpha, and Hoyt's method (Rivera, 2007). In validity, the classical test theory assesses the predictive validity, concurrent validity, content validity, construct validity and face validity. Hence for determination of item effectiveness, validity and reliability of the developed BPST, classical test theory (CTT) approach formed the conceptual framework.

Theoretical framework in the second stage of the study

1.9.3 Objectives-Oriented Approach to Evaluation (Tyler, 1949)

The objectives-oriented approach to evaluation by Tyler (1949) formed the conceptual framework for this study in the second stage where the study assessed the general performance of Biology students in the science process skills test (BPST). The Study also assessed the performance of students in each process skills under study (formulating hypotheses, defining operationally, controlling variables, planning investigations as well as interpreting data). Objective-based approach by Tyler (1949) defines evaluation as a process of determining the degree to which educational objectives are being achieved. Of all the approaches to curriculum evaluation, the objectives-oriented approach is arguably the simplest and most straightforward to use (Luo et al. 2005). It follows the scientific tradition and is straightforward to apply, but does not take account of unintended outcomes, and takes no account of students as individuals with all their differences (Luo et al. 2005). The focus of an objectives-oriented evaluation is on specifying goals and objectives and determining the extent to which those goals and objectives have been attained by the program in question (Luo et al. 2005). Tyler's scheme set up seven specific steps to conducting an educational evaluation based on objectives; it was clearly influenced by behaviorism, which was dominated psychological thought at the time. The steps includes i. establish broad goals/objectives ii. classify them in an orderly manner, iii. define the objectives in behavioral terms, iv. find situations in which achievement of the objectives can be concretely shown, v. develop or select appropriate measurements, vi. collect (student) performance data, and vii. compare the performance or outcome data with the objectives set before. Tyler stressed that a wide range of objective and performance assessment procedures usually should be employed. Criterion-referenced tests and students' work samples are especially relevant to this evaluation approach (Tyler, 1949)

Theoretical framework in the third stage of the study

In investigating the effectiveness of inquiry-based approach on students' scientific process skills development, conceptual understanding of Biology contents and motivation, this study was guided by three theories/models. These theories included i. 5E inquiry-based instructional model by Bybee et al. (2006), ii. constructivism learning theory, and iii. the social-cognitive motivation theory (1986, 1997).

1.9.4 Bybee, et al. (2006) 5-E inquiry instructional model to science (learning cycle model)

Bybee et al. (2006) 5-E inquiry instructional model to science guided the investigation of the effectiveness of inquiry-based approach on students' scientific process skills development, conceptual understanding and motivation. The 5 E's is an instructional model based on the constructivist approach to learning, which emphasizes on the need of learners to build or construct new ideas on top of their old ideas (Bybee et al. 2006). Inquiry-based teaching was defined in this study as the product of the blended theories of Piaget, Vygotsky, and Ausubel about the philosophical underpinnings of teaching and learning known as constructivism (Liang & Gabel, 2005). Constructivism emphasizes the active thinking process of integrating prior knowledge with existing knowledge (Bybee et al. 2006; Bybee, 2012). The thinking went further to overcome one big misconception that, oftentimes inquiry is equated to the scientific method. However, in this study, the inquiry was not restricted to scientific procedures. As stated by the AAAS (1993) that inquiry is not restricted to the following of the steps of the scientific method only. In as much as most science activities are presented with procedures, teachers should bear in mind that there is more to it when they are doing inquiry (AAAS, 1993). These procedures must only serve as an awakening statement in order for the students to formulate their own questions as teachers guide and mentor them in the process (AAAS, 1993).

Each of the 5 E's describes a phase of learning, and each phase begins with the letter "E": Engage, Explore, Explain, Elaborate, and Evaluate. The 5 E's allows students and teachers to experience common activities, to use and build on prior knowledge and experience, to construct meaning, and to continually assess their understanding of a concept (Bybee et al. 2006). The use of this model brings coherence to different teaching strategies, provides connections among educational activities, and helps science teachers make decisions about interactions with students (Bybee et al. 2006). Teachers can use the 5-E model as seen in figure 1.1 below, to meet objectives and deliver specific concepts and explanations.

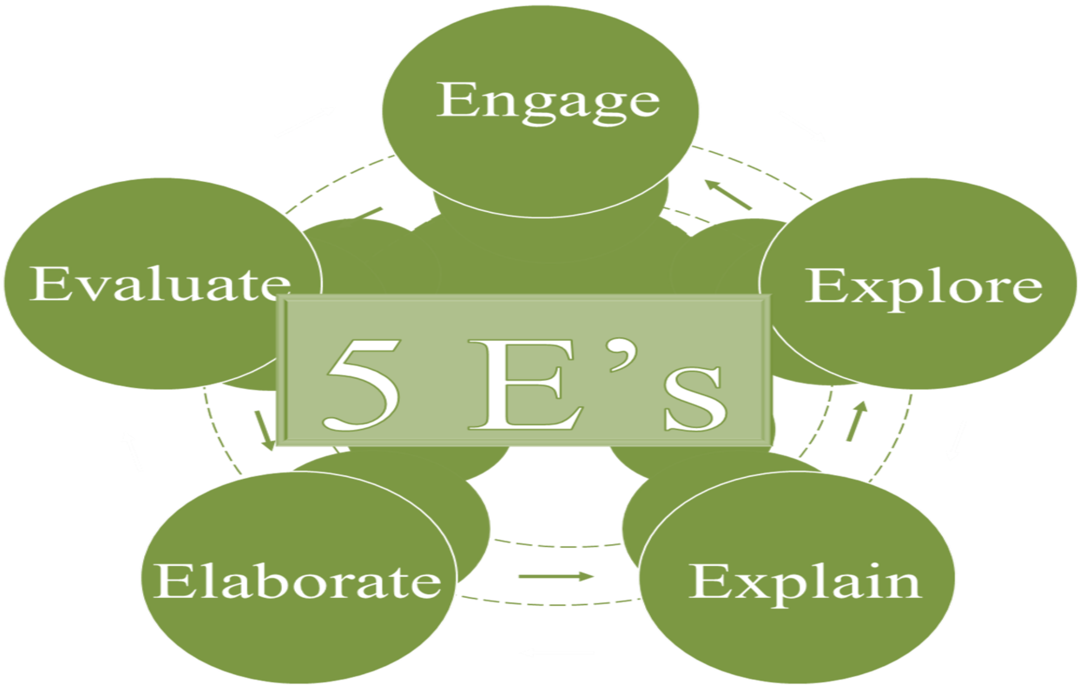


Figure 1.1 Learning cycle modified from Bybee et al. (2006)

This model follows a step-by-step progression, (see figure 1.1 above) where each step builds on the previous step. In the engagement phase, the task is introduced by making connections to past learning and experience through demonstration of an event, the presentation of a phenomenon or problem or asking pointed questions can be used to focus the learners' attention on the tasks that will follow (Bybee et al. 2006). The goal is to spark their interest and involvement. In the next phase (Explore) learners take part in activities that allow them to work with materials that give them a 'hands on' experience of the phenomena being observed. Learners can be provided with simulations or models whose parameter can be manipulated by learners so that they can build relevant experiences of the phenomena (Bybee et al. 2006). Questioning, sharing, and communication with other learners are encouraged during this stage. The teacher's role is only to facilitate the process. In explaining phase, the focus at this is on analysis. The learner is encouraged to put observations, questions, hypotheses and experiences from the previous stages into the language (Bybee et al. 2006). Communication between learners and learner groups can spur the process. The instructor may choose to introduce explanations, definitions, mediate discussions or facilitate by helping learners find the words needed. The fourth phase is elaboration students are supposed to use and build understanding gained in the previous stages to expand upon it (Bybee et al. 2006). Lastly, evaluation should be an ongoing process and should occur at all stages, in order to determine whether learning objectives are met and misconceptions avoided. Any number of rubrics, checklists, interviews, observation or other evaluation tools can be used. If interest in a particular aspect or concept is shown, further inquiry should be encouraged and a new cycle can begin that builds upon the previous one (Bybee et al. 2006).

In this study, genetics teaching modules of the experimental group students were modified and implemented in the manner that it gives students opportunities to engage, explore, explain, elaborate, and evaluate the learnt information. Contrary from the control group, the lessons of the experimental group students begin with an activity, a challenge, or a specific question to discuss, or an article for reading to allow for 5E sequence to occur.

1.9.5 Constructivism and behaviorism theories of learning

In connection with the inquiry-based 5E instructional model, constructivism and behaviorism theories of learning were other theoretical underpinnings which laid the foundation for this study in the third stage. As it has been stated, the stage intended to investigate the effectiveness of inquiry-based approach on students' scientific process skills development, conceptual understanding of genetics topic and motivation towards science process skills by comparing it with the conventional approach. It has to be noted that the competence-based curriculum of Tanzania of 2005 recommended that science instruction and learning should be well grounded in inquiry and teachers should move from conventional approaches to participatory based method.

According to the proponents of constructivism, knowledge is constructed as students integrate new information with their pre-existing knowledge base (Eiskenkraft, 2003; Ramsey, 1993). Educational movements such as inquiry-based learning, active learning, experiential learning, and discovery learning are variations of constructivism. The theory suggests that students learn science best when they are actively engaged in doing science or in performing activities that allow them to think like scientists (Burris & Garton, 2007; Doolittlen & Camp, 1999). As such, a major emphasis in science curricular reform is a change from a more traditional teacher-based learning to a more inquiry-based, student-centered learning (Burris & Garton, 2007). The role of a constructivist teacher is to set up problems and monitors student exploration and discussions, guides student inquiry, and promotes new patterns of thinking. Because of the emphasis on students as active learners, constructivist strategies are often called student-centered instruction (Eiskenkraft, 2003). Teachers are supposed to ask questions to explore learner's previously constructed information – looking for preconceptions (Glasersfeld, 1995). They then lead learners through exploratory activities that enable them to investigate on their own and come to their own conclusions (Brooks & Brooks, 1999). Student-centered constructivist teaching shift the focus of activity from the teacher to the learners in which students solve problems, answer questions, formulate questions of their own, discuss, explain, debate, or brainstorm during class (Rolf et al. 2010; Eiskenkraft, 2003)

From a constructivist point of view, the well-known didactic triangle for interactive constructivist teaching (Rolf et al. 2010) is presented in figure 1.2 below.

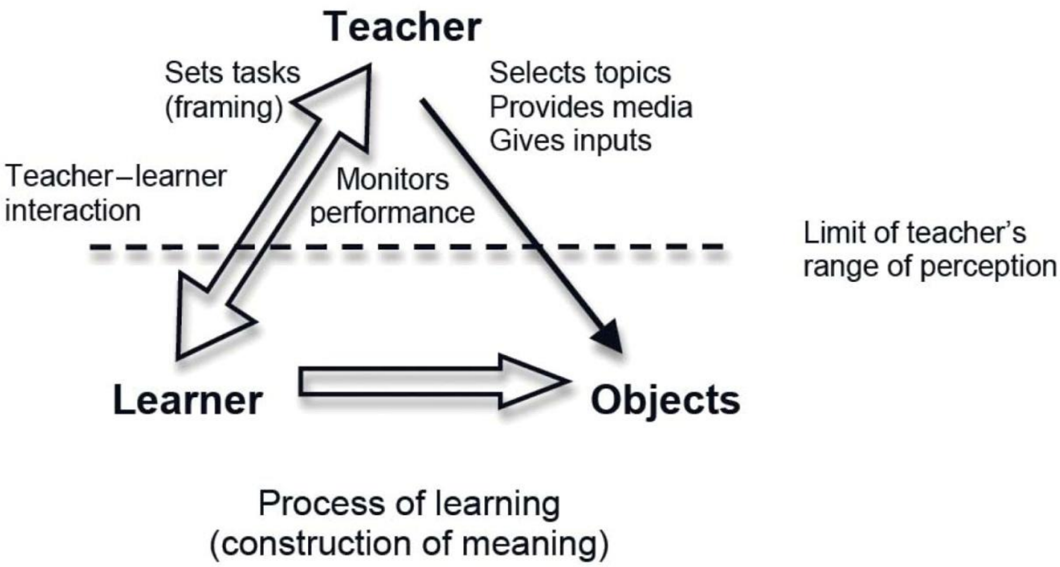


Figure 1.3 Interactive constructivist teaching (Rolf et al. 2010)

In triangular above (figure 1.2), it is the learner who creates his or her understanding of the objects of learning. The construction of meaning takes place in the learner's mind, beyond the teacher's range of perception. What the teacher sees is the outcome – what students produce, and how they behave (Rolf et al. 2010). Teachers can support their students by creating learning opportunities, designing challenging tasks, providing instruction through media and inputs (lectures) that represent the objects of learning and encouragement and support for self-esteem (Rolf et al. 2010). Learners reconstruct what they have learnt and they apply it and put it to the test., the teacher provides them by giving opportunities for sharing, presentation and discussion, formal testing and assessment and by designing challenging tasks (Rolf et al. 2010).

In this study, constructivism teaching approaches which are participatory and more interactive methods were applied in the experimental group in teaching themes within genetics. The role of the teacher (the researcher) was to promote discussion, active learning, and reflection, and provide modeling, coaching, and scaffolding to students when required. The teacher (the researcher) acted as a facilitator rather the custodian of knowledge. Learning activities to the experimental group were designed in such that they require full student engagement in practical, real world tasks and allow opportunities for students to reflect on their learning experiences. The researcher just offered a platform where the learners interacted with sources of knowledge reconstructs knowledge and takes responsibility for their learning.

On the other hand, the teaching of genetics to the control group students was guided by behaviorism theory of learning. Behaviorism is a learning theory that assumes a learner is essentially passive, a blank slate (Tabula rasa) responding to environmental stimuli (Chung, 2004). According to Skinner (1974), behaviorism theory focuses only on objectively observable behaviors and discounts any independent activities of the mind. It views behavior as the product of conditioning. The learner starts off as a clean slate (i.e. tabula rasa) and behavior is shaped through positive reinforcement or negative reinforcement (Barakatas, 2005). According to Skinner (1974) originators and important contributors to this theory includes John B. Watson (classical conditioning), Ivan Pavlov (classical conditioning), B.F. Skinner, E. L (operant conditioning) and Thorndike, E (connectionism). Behavior theorists define learning as nothing more than the acquisition of new behavior based on environmental conditions (Chung, 2004). The theory treats humans as biological machines and do not consciously act, rather they react to stimuli (Floresc & Kaylor, 2007).

Teacher's role in the behaviorist class (traditional paradigm) is to decide, control and direct student learning (banking education) (Barakatas, 2005). The teacher is an authority who decides what and how their students should be learning (Chung, 2004). Teaching practice is based on the assumption that learning is a mental process which substitutes one stimulus for another in conditioned responses i.e conditioned learning or conditioning (Chung, 2004). Lessons are designed with a view to specific learning outcomes which are outlined in structured lesson plans. Evaluation of learning is based on student performance on objective tests (Floresc & Kaylor, 2007).The theory has been criticized as it emphasizes on conditioning and rote learning involves the unnatural imposition of meaningless stimuli on the brain(Chung, 2004). In the behaviorist view, knowing is an organized accumulation of associations and components of skills. Learning is the process in which association and skills are acquired (Chung, 2004).

In this project, this mode of teaching and learning was termed as a traditional or conventional method and it referred to as the teacher directed learning style that can also be called behaviorist. In some cases, this behaviorist approach can also be called transmission, instructivism, teacher-centered, direct instruction, static view or mechanistic view (Barakatas, 2005). This approach was employed in teaching the control group students. Lecture notes and discussion questions were prepared in advance before the actual class session. Three different textbooks prescribed by the Tanzania Biology syllabus and proved adequate to provide the essential factual basis for the course and were used in the construction of student's notes and discussion questions.

1.9.6 Social-cognitive theory on motivation

The conceptual framework on motivation in this study was based on a social-cognitive framework postulated by Bandura (1986, 1997 & 2001) where the motivation to learn science process skills is conceptualized to having both cognitive and affective influences. Within this framework, motivation is defined as the internal state that arouses, directs, and sustains students' behavior toward achieving certain goals (Broussard & Garrison, 2004). The social-cognitive theory attempts to explain why students strive for particular goals, how intensively they strive, how long they strive, and what feelings and emotions characterize the process. In this case, individuals' thoughts, beliefs, and emotions are central processes that underlie motivation (Glynn et al. 2006). Hence motivation to learn is defined as a student's tendency to find academic activities meaningful and worthwhile and to try to derive the intended academic benefits from them (Broussard & Garrison, 2004). Within social-cognitive theory, each student is viewed as possessing a self-regulating system that affects beliefs and aids in the development of motivation that enables behavior cognitively and affectively (Schunk & Pajares, 2001). A central tenet of Bandura's (1986, 1997, 2001) social cognitive theory is that human behavior operates within a framework of triadic reciprocality involving reciprocal interactions among three sets of influences: personal (e.g., cognitions, beliefs, skills, affect); behavioral; and social/environmental factors (Schunk, & Usher, 2012).

Social cognitive theory assigns a prominent role to self-regulatory processes (Schunk, & Usher, 2012; Bandura, 2001). Self-regulation refers to the processes that individuals use to personally activate and sustain behaviors, cognitions, and affects, which are systematically oriented toward the attainment of goals (Schunk, & Usher, 2012; Zimmerman, 2000). In the context of learning, prior to embarking on learning the process, students set goals and determine which strategies to use. They then regulate their behaviors to conform to their internal standards and goals. As they work on tasks, they assess their progress toward their goals and decide whether to continue or alter their strategies (Schunk, & Usher, 2012). During breaks or when tasks are completed, they reflect on their experiences, seeking to make sense of them and to determine what their next steps should be (Schunk, & Usher, 2012). As they reflect on what they have done, their beliefs that they have learned strengthen their self-efficacy and motivate them to continue learning (Schunk, & Usher, 2012). Thus the self-regulatory system affects a student's academic achievement by influencing behaviors such as class attendance, class participation, question asking, advice seeking, studying, and participation in study groups (Pajares & Schunk, 2001; Glynn et al. 2006).

There are at least five key constructs within the self-regulatory system that would contribute to a student's overall motivation to learn science process skills and, consequently, achievement (Glynn et al. 2006; Glynn et al. 2011). The first one is intrinsic motivation, the kind of motivation to perform a task for its own sake (Ryan & Deci, 2000). It is the self-desire to seek out new things and new challenges and it is driven by an interest or enjoyment in the task itself, and exists within the individual rather than relying on external pressures or a desire for reward (Ryan & Deci, 2000). The second one is goal orientation motivation or extrinsic motivation. In this case students with learning goals tend to be motivated in learning subjects like science as a means to a tangible end such as a career, grade, recognition, rewards, selection etc (Ryan & Deci, 2000). Extrinsic motivation comes from influences outside of the individual. However, both types of motivation, are (intrinsic and extrinsic) important in contributing to students' success in their courses (Ryan & Deci, 2000). However, extrinsic motivation according to Glynn et al. (2011) may be sub divided into two types of motivations (the grade motivation and career motivation). Grade motivation and career motivation target more precisely the primary ends that student focus on (Lin et al. 2003). Grades are important short-term goals because they are measures of school success and part of the entry criteria for many careers (Glynn et al. 2011). Careers are important long-term goals (Glynn et al. 2011). The fourth construct is self-determination where students believe they have some degree of control over their learning such as selecting some of their topics (Ryan & Deci, 2000; Reeve et al. 2003). When students have the opportunity to choose what their assignments will be, they are more likely to learn from the assignments than vice versa (Glynn & Koballa, 2006).

The fifth construct of motivation in the context of the socio-cognitive theory is self-efficacy which is defined by Bandura (1997) as beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments. Bandura's self-efficacy theory classifies people's behavior into two major categories; people with a strong self-efficacy and those with a weak sense of self-efficacy (doubting their capabilities in difficult situations). People with a strong sense of self-efficacy according to Bandura (1977), tend to approach difficult tasks (learning of scientific skills in this case) as challenges to be mastered with assurance in themselves about their capabilities. This type of outlook is seen to produce personal accomplishments, reduce stress, and lower vulnerability to depression. On the other hand, people with low self-efficacy tend to have low aspirations and weak commitment to the goals they pursue. They easily develop stress and depression which in turn, hamper their capacity to perform actions effectively (Bandura, 1977).

However related to self-efficacy and self-esteem is the concept of self-concept. Self-concept generally refers to the totality of a complex, organized and dynamic system of learned beliefs, attitudes, and opinions that each person holds to be true about his or her personal existence. Byrne, (1990) has defined global self-concept as the way (positive or negative) people feel about themselves in general. A person sees himself as a successor or a failed only in relation to his experience with others or in the way those experiences having been interpreted for him. Therefore motivation of students towards science process skills in this case partly depends on the perceived self-concept of an individual student. In light of social-cognitive conceptualizations, motivation towards science process skills in this study was conceptualized as being influenced by their intrinsic motivation, perceived self-efficacy, career motivation, perceived self-determination, grade motivation, and perceived self-concept.

1.10 Significance of the study

i. In the early 2000s, Tanzania began a process of curriculum reform with the goal of transforming Tanzania schooling from exam-oriented education to student-centered learning. Traditional education practices had expected students to passively accept and memorize material presented by teachers, and to reproduce the knowledge on often high-stakes examinations. As a result of these transformations, in 2005 Tanzania came up with the so called 'Competence Based Curriculum' which emphasized among other things, student's competence in science process skills. The curriculum emphasized the need of Tanzania science students to learn scientific subjects such as Biology, Physics and Chemistry in the same way science is done scientists. Despite such a dramatic shift in curriculum policy, little been known about whether or not the reform efforts are truly transforming the educational experiences of students. One of the significance of this study is the fact that it has established a base level of information on the knowledge level of Tanzania science students in the area of science process skills.

ii. The study firstly developed and validated a science process skills test specific for Biology which will be used in assessing students' competence in this area. It is hoped that the developed test will provide educators with a valid, reliable and cost effective means of measuring science process skills efficiently and objectively. Furthermore, the developed test will provide a practical solution of assessing the science process skills in large under-resourced science classes.

iii. This research is also significant to the domain of science education as it extends the knowledge base with respect to the concept of inquiry and constructivist approach to science teaching. The study compared an activity-based, inquiry-oriented approach using simple, inexpensive, available instructional materials with a traditional teaching approach I the achievement and understanding of the concepts related to genetics. Hence the study aims to obtain information on which one could make an informed decision about the relative suitability of the two teaching approaches.

iv. The findings of this study could also encourage other scholars to investigate the manner in which other competences addressed in the new 2005 education curriculum, apart from science process skills could be taught, or integrated in teaching.

1.11 Limitations of the study

i. One of the major limitations of this study stems from the fact that it required a large amount of empirical data if generalization of the findings was to be established. Owing to time and resources constraints, it was not possible to collect data from all advanced level Biology students in secondary schools across the country. Although it was realized that such a collection would greatly enhance external validity of the findings.

ii. For students to demonstrate the integrated process skills, assessment using hands-on procedures (authentic assessment) to determine skill acquisition by students is pivotal. The use of a paper and pencil test to assess practical skills like BPST has been criticized by several researchers who advocate for practical manipulation of apparatus and physical demonstration of practical skills. This presents a limitation in the findings of this study in the sense that the instrument used in this study (BPST) does to accommodate these requirements. (More specific limitations and constrains for each stage of the study have been intensively discussed in chapter seven of this report).

1.12 Delimitation of the study

This study was conducted within the following parameters to delimit its scope;

i. Only advanced level secondary school Biology students in Morogoro municipality were involved in the study. This delimited the overall magnitude of the study.

ii. Science process skills are classified into, (a) Basic and (b) Integrated skills. This study dealt only with assessing students' ability in integrated science process skills. This is because integrated skills are higher order and complex than basic skills. Dealing with integrated science process skills only also delimited the scope of study.

CHAPTER TWO

LITERATURE REVIEW

2.1: Introduction

This chapter is about the review of literature and studies related to test development and validation, science process skills, inquiry-based approach to science teaching and motivation in science education. The chapter comprises of summaries of how different scholars approached the topics of science process skills, test development and validation, inquiry approach to science and their major findings. The review is organized under the following major sub-headings; i. the meaning and importance of science process skills to students, ii. history and teaching of science process skills in Tanzania science curricular iii. description of individual integrated science process skills, iv measurement of science process skills, v. criteria and procedures for science process skills' test development and validation, vi. the meaning and importance of inquiry-based instruction in the development of students´ science process skills, conceptual understanding of contents vii, motivational components and their role in science learning, and lastly the literature gap.

2.2 Literature on the meaning and importance of science process skills

2.2.1 Meaning and classification of science process skills

There are a number of ways of conceiving and categorizing the science process skills. Several definitions by different science educationists exist which describes the meaning of science process skills (Chiappetta & Koballa, 2002; Padilla et al. 1983; Westbrook & Rogers, 1994; Pratt & Hackett, 1998; Rezba et al. 1995). According to Rezba et al. (1995) science process skills are a means for learning and are essential to the conduct of science. Nwosu & Okeke (1995) described science process skills as mental and physical abilities and competencies which serve as tools needed for the effective study of science and technology as well as problem-solving. Chiappetta & Koballa (2002) on the other hand define science process skills as a set of broadly transferable abilities appropriate to many science disciplines and reflective of the behavior of scientists. These skills are appropriate for all science fields, and they reflect on the correct behaviors of scientists while they are solving a problem and planning an experiment (Padilla, 1990).

As with the meaning, there is also no a universally common classification of science process skills. The American Association for the Advancement of Science (AAAS) classified the science process skills into fifteen; observing, measuring, classifying, communicating, predicting, inferring, using number, using space/time relationship, questioning, controlling variables, hypothesizing, defining operationally, formulating models, designing experiment and interpreting data (Padilla,1990; Espinosa et al. 2013). Herlen (1999) on the other hand has classified science process skills into observing, hypothesizing, predicting, raising questions, investigating, interpreting, communicating, respect for evidence, flexibility, and critical reflection. Zimmerman (2000) classified science process skills as specific for a field, or general process skills, and also argued that knowing the scientific terms of the issue must be achieved in order to solve any problems about one issue.

However, most scientists tend to classify science process into two categories; the basic and integrated process skills (Chiappetta & Koballa, 2002; Rezba et al. 1995; Mungandi, 2005; Brotherton & Preece, 1995; Hamilton & Swortzel, 2007). The basic (simpler) process skills provide a foundation for learning the integrated (more complex) skills. Six basic science process skills according to Brotherton & Preece, (1995) include observation, communication, classification, measurement, inference and prediction. Chiappetta & Koballa (2002 listed the integrated science process skills as skills in identifying and controlling variables, defining variables operationally, formulating hypotheses, interpreting data and in designing experiments. The current study has adopted the above basic vs integrated classification system of science process skills where the focus will be on the higher order skills called integrated science process skills (see section 1.9.1).

2.2.2 General education importance of science process skills

One of the most important and pervasive goals of schooling is to teach students to think. Numerous researchers have written on the importance of science process skills in helping students to think scientifically and increase their conceptual understanding. According to Rezba et al. (1995) students who have learned scientific process skills have the tools to interpret what they observe, make inferences and predictions about their observations. Rehorek (2004) on the other hand believe that development of science process skills enables students to construct and solve problems, critical thinking, deciding and finding answers to their curiosity, rather than having the students to memorize the concepts. Besides being the thinking skills that students can use to get information, science process skills according to Bredderman (1983), also guides students think scientifically on the problems and formulate testable hypothesis and experiments. It is more important for the students to learn how to apply science than learning reality, concepts, generalizations, theories and laws in science lessons. Some researchers have reported that science process skills practical works increase students' sense of ownership of their learning and can increase their motivation (Johnstone & Al-Shuaili, 2001). Mungandi (2005) supported this perception by arguing that, if learners have to be true future scientists, they need to learn the values and methods of science process skills. It is through process skills that scientific ideas are developed, tested and linked. Process skills are not only important as part of the core skills in sciences, but also in enabling learners to develop the ability to use evidence in solving problems and making decisions (Harlen, 2000).

Motivation and positive belief of students are significant components for effective learning in science (Taylor & Corrigan, 2005). In Fishbein's model, beliefs affect attitudes and these attitudes then affect intentions and behaviors (Weinburgh & Englehard, 1994). Science process skills also have motivational significance to science learners. Numerous literature exists describing how science process skills affect students' interest motivation and self-belief toward science (Tuan et al. 2005; Wolters & Rosenthal, 2000; Rowland, 1990). According to Arena (1996) students are more motivated when they can test their ideas that have relevance to them in the lab inquiry style experiments. Thompson & Soyibo (2001), in a comparison study, reported positive impacts of a combination of demonstrations, discussion and practical work focusing on science process skills on Jamaican 10th grade [age 15-16] students' attitudes. Other researchers claimed that science process skills increases students' attitudes and interest in science and scientists (Lunetta et al. 2007), encourages students to acceptance of scientific inquiry as a way of thought(Harlen, 1999; Wynne 1999), enables students´ adoption of scientific attitudes (Rowland, 1990), increase students enjoyment of science learning experiences(Harlen, 1999; Wynne 1999), develops students´ interest in science and science-related activities (Kok-Quntoh & Woolnough, 1994), and develops of interest in pursuing a career in science (Lunetta et al. 2007).

Some researchers (Beaumont-Walters, 2001; Souchek & Meier, 1997; Mungandi, 2005; Padilla, 1990; Preece & Brotherton, 1997; Harlen, 2000; Kok-Quntoh & Woolnough, 1994) have described science process skills as special skills that simplify learning science, activate students, develop students' sense of responsibility in their own learning and increase the permanency of learning. In supporting this contention Chiappetta & Koballa (2002) wrote that the acquisition and frequent use of these skills can better equip students to solve problems, learn on their own, and appreciate science. In the same line of thinking, Mungandi (2005) commented that science process skills are the instruments that scientists use to learn about the world and empower learners with the ability and confidence to solve problems. Beaumont-Walters (2001) on the other hand wrote that hands-on and learning by experience are powerful ideas, and we know that engaging students actively and thoughtfully in their studies pay off in better learning.

2.3 Science process skills from the experience of Tanzania education system

2.3.1 The history of science process skills in the Tanzanian science curricular

The independent Tanganyika 1961 inherited an education system and a curriculum available only to a minority and distinguished by race, economic position, geographical location, and religious denomination (Galabawa, 2005; Ishumi & Nyirenda, 2002). Curricula, syllabuses, examinations, textbooks, teaching and learning materials were inferior, less equipped, and irrelevant to Africans as it was based on British prototypes (Sefu & Siwale, 1977). The incorporation of science process skills in Tanzanian education curricula dates back to 1967 when the government introduced Education for Self-Reliance (ESR) policy which re-defined the purpose of education in the country. According to Osaki (2007) the redefinition of education meant to develop learners' inquiry mind, critical thinking, confidence, and mental liberation. The philosophy of education for self-reliance placed much emphasis on practical activities whereby classroom works was to be linked with real life. As a result, science classroom teaching methods were transformed to emphasize experimentation and actual experience (Ishumi & Nyirenda, 2004; Sefu & Siwale, 1977). This orientation, in some ways, provided students with opportunities to practice what they have learned in the classroom, test their ideas and analyze data or information observed thereby developing science process skills.

A typical curriculum based on science process skills according to Osaki (2007), was firstly introduced in Tanzania in 1968. The curriculum was an effort by the School Science Project (SSP) of 1968 which sought to change the way science was taught from a theoretical mode to an inquiry manner where pupils would learn to think and solve authentic scientific problems. It was an activity and inquiry based curriculum and it was adopted from Nuffield science materials (Osaki, 2007). The curriculum covered all science subjects and focused on an experimental approach that also touched the historical development of ideas in each topic and a great deal of outdoor and laboratory activities (Osaki, 2007). As a result, Biology learning started to involve a lot of ecological sampling, collecting and identification of specimen and doing experimental write-ups. Chemistry lessons had a lot of practical works such as analysis of substances and titrations. Physics, on the other hand, had lots of field visits, measurement taking, calculations, and games all of which contributed to students' science process skills (Osaki, 2007).

However, Osaki (2007) reported that the SSP inquiry curriculum was abandoned in the mid-1970s following a mass failure of the experimental group in their final national examinations as compared to the control group which was conventionally taught. Students' failure was translated as inquiry curriculum weakness. However, according to (Osaki, 2007), students' failure was a result of insufficient teacher orientation on the requirements of the new curriculum. The Institute of Education, therefore, abandoned the emphasis on inquiry science and continued to write textbooks that focused more on remembering facts and formulae instead of scientific experimental works (Osaki, 2007)

2.3.2 Teaching of science and science process skills in Tanzania

For many years, science and hence science process skills teaching in Tanzania has been poorly and traditionally conducted with a focus on content knowledge than the scientific skills (see Semali & Mehta 2012; Athuman, 2010; Chonjo et al. 1996; Mushi, 1992; Osaki, 2007). For example, the study conducted by Athuman (2010) on the teaching of science process skills in Morogoro found domination of teacher-talk lecture method in many science classes. The researcher found science teachers confining students to observing and listening passively without engaging in activities as how science lessons should be like (Athuman, 2010). According to Athuman (2010), lack of references and overcrowding of students in the classrooms exacerbated the problem of teaching by lecture method only without scientific activities. In another study that was conducted in the middle of 1990s by Chonjo et al. (1996) to investigate the status of science teaching in Tanzanian secondary schools revealed that there were deep-rooted problems in the teaching and learning of science in the secondary education sub-sector. According to Chonjo et al. (1996), science teaching in Tanzania is challenged by weak teacher pedagogical competence, overloaded curriculum, much emphasis on examinations, ill-equipped laboratories, and low teacher motivation.

The above findings resemble those by Semali and Mehta (2012) on their study titled "*Science education in Tanzania: Challenges and policy responses*" which revealed that both teachers' pedagogical skills and practical skills are inadequate, and they continue to lecture, with a focus on the next examinations rather than promoting understanding. Likewise, Semali & Mehta (2012) found science syllabuses being overloaded with content and examination requirements compel teachers to teach in a rush to cover the long syllabus. Subsequent studies (Chonjo & Welford, 2001; Kibga, 2004; Leeuw, 2003; Mafumiko, 1998) showed that the situation in most schools did not significantly improve over time, especially in relation to resources and the conduct of practical work, shortage of science and mathematics teachers, as well as shortage of textbooks.

Laboratory experience has been given a central and distinctive role in science and educators have listed a number of benefits of teaching by using laboratory activities (see Lunneta & Hofsin, 1982). In Tanzania, however, this seems not the case. For example a study by Mushi (1992) on the teaching of Physics in some selected secondary schools in Tanzania found that in Form III, out of a total of 78 periods, only 18 were assigned for laboratory activities necessary for science process skills, and in Form IV only 2 out of the 63 periods were designated for this purpose. There was no opportunity for students to use meta-cognition and constructivist thoughts rather they were expected to absorb the teacher's constructed meanings of the lessons (Mushi, 1992). Similar findings were also witnessed by Osaki (2007), who noted that only few laboratory works for scientific skills were assigned to students after instruction and they were designed to confirm the lesson rather than investigate a phenomenon. With this kind of teaching, a lot of questions emerge as to whether students would really acquire scientific skills.

According to Osaki (2007) this is because science teaching in Tanzania has been driven by the pressure of performance on high-stakes testing thus over emphasizing content knowledge than laboratory activities which basically equip learners with science process skills. The author further argues that laboratory practical for science process skills in Tanzania had virtually been stopped due to heavy financial costs involved and little government budget support. A weak government support forced the National Examination Council of Tanzania (NECTA) in the mid-1980's to abandon real practical examinations of science subjects. Instead, a costless and theoretical alternative to practical examinations were introduced. This shift was interpreted by science teachers to mean that there is no need for practicals and hence scientific skills (Osaki, 2007).

There is a general consensus among science education researchers inquiry-based approach should be an integral part of science teaching if students are to acquire science process skills (Prince & Felder, 2006; Haury, 1993; Rubin & Norman,1992; Shymansky, 1990; Crawford, 2000). However, a survey conducted by Chonjo et al. (1995) on the teaching of science in selected schools in Tanzania found that very few teachers were committed to using inquiry approaches. Teaching sessions were boring and teachers did not use a variety of teaching strategies. Chonjo et al. (1995) further found that science teaching had been reduced to copying and memorization of facts for examinations. Osaki & Njabili (2004) conducted another survey on the teaching of science in secondary schools and observed that learners were being put into groups in the name of participatory teaching with many of them looking bored and confused as to what were points taught by the teachers, and some obvious wharfing. Teachers lectured in a didactic fashion, droning and carrying on until the end of the class (Osaki & Njabili, 2004)

In schools, science process skills can be well taught and developed during laboratory, research and other investigative activities, project works and in various inquiry learning experiences (Dyer and Myers, 2006). According to Roth and Roychoudhury (1993), laboratory experiences provide students with the freedom to perform experiments of personal relevance in authentic context. Through experiments students learn to, (a) identify and define pertinent variables, (b) formulate testable hypotheses to guide investigations, (c) plan and design experiments, and (d) analyze, transform, and interpret data (Roth & Roychoudhury, 1993). These are essential and basic activities which are necessary for students' acquisition of science skills.

Mechling et al. (1985) accord that, if science process skills have to be learned well, they should be combined with science contents enabling students to learn them with contents at the same time in a seamless learning experience. The authors further argued that in teaching science process skills whether in a classroom or in the laboratory, there must be competence indicators which have to be spelled out clearly as to what students should be able to do to achieve a mastery of a scientific skill. Teachers guide student learning by selecting, designing and planning learning tasks, asking probing questions, observing students at work to identify misconceptions and planning follow-up experiences.

2.4 Description of individual integrated science process skills under study

Several writers have attempted to describe individual integrated science process skills and the context in which they may be developed and imparted to learners (Thompson & Soyibo, 2001; Wetzel, 2008; Padilla, 1990; Rambuda, & Fraser, 2006; Osborne et al. 2003; Chiappetta & Koballa, 2002; Harlen, 2000; and Padilla et al. 1983; Preece & Brotherton, 1997; Walters & Soyibo, 2001; Minstrell & van Zee, 2000). This section make discussion and descriptions of the five integrated science process skill under the current study i. Hypothesis formulation ii. identifying and controlling variables, iii. defining operationally iv. analysis and interpretation of data, and v. experimenting in the following subheadings.

2.4.1 Hypothesis formulation skill

The scientific method requires that one should test a scientific hypothesis. A hypothesis is an educated guess that can be scientific verified. Wetzel (2008) defined a hypothesis as a process of making a prediction (intelligent guess) based on evidence of prior investigations or the expected outcome of an experiment. There are a number of literature and studies on the centrality of hypothesis in scientific formal thinking (Gay & Airasian, 2000; Wetzel, 2008; Herlen, 1999; Seok Oh, 2010; Ary et al. 1990). According to Gay & Airasian (2000), it is imperative to formulate a testable hypothesis which directs the way investigation should be designed and eventually carried out. Hypotheses predict about the relationships between variables and guide the researcher with regard to the kinds of data to gather. According to Rezba et al. (1995), prediction should be based on facts, opinion, hunch, or whatever resources one may possess. Scientists generally base such hypotheses on previous observations or on extensions of scientific theories.

As one of the core integrated science process skills, several studies on how teachers could best develop hypothesis formulation skill to learners have been conducted. Seok (2010) conducted a study to find out how the teacher could help students formulate scientific hypotheses. His analysis identified four categories of teaching strategies which could be used by science teachers to help students in hypothesis-generating inquiry. These included: (1) expanding and activating students' background knowledge, (2) providing analogies, (3) questioning, and (4) encouraging students to use alternative forms of representation (Seok, 2010). According to Harlen (2000), when hypothesizing, the suggested explanation need not be correct, but it should be reasonable in terms of evidence and be possible in terms of scientific concepts or principles. The author added that at early stages of developing hypothesis formulation skill, a learner is expected to make an attempt to explain something based on his/her earlier experience. Teachers should make sure to follow each step of the scientific method with each experiment and get students to be as involved as possible with helping to formulate a hypothesis and help to solve the problem or question at hand (Harlen, 2000). This encourages students to formulate their own personal opinions and come to their own conclusions.

2.4.2: Identifying and controlling variables skill

A variable is anything that affects the results of your experiment. In scientific inquiries, scientists investigate to find out what causes something to happen, that means to find the effect of one variable on another. Identifying and controlling variables according to Padilla et al. (1983) is the ability to identify variables that can affect an experimental outcome, keeping variables constant while manipulating only the independent variables. The skill involves manipulating and controlling properties that relate to situations or events for the purpose of determining causation (Chiappetta & Koballa, 2002).This means firstly, identifying the variables in a situation, then selecting variables to be manipulated and held constant (Chiappetta & Koballa, 2002).

Numerous scientific research method books have written on the meaning, classification and importance of identifying and controlling experiment variables (Bauer, 1992; Gauch, 2003; Ziman, 2000; Ross, 1988). According to Ross (1990), there are three kinds of variables in a scientific investigation: independent, dependent, and controlled. The independent variable is the one that is changed by the scientist. To ensure a fair test, a good experiment has only one independent variable (Ross, 1990; Miller, 2003; Bauer, 1992). The scientist focuses his or her observations on the dependent variable to see how it responds to the change made to the independent variable. Bauer, (1992) defined the dependent variable as the variable that is measured by the experimenter. Experiments also have controlled variables. Borg & Gall (1989) defined controlled variables as quantities that a scientist wants to remain constant, and he must observe them as carefully as the dependent variables. They are the variables that are controlled or more specifically kept constant so that they do not unduly affect how the independent variable affects the dependent variable (Ross, 1990; Ross, 1988; Ziman, 2000). In practical investigations, the experimental group is usually exposed to some treatment while the control group is not exposed to such treatments (Miller, 2000; Borg & Gall, 1989).

The idea behind identifying and controlling variables is that we must be sure that what we think caused an effect did in fact cause it. This skill requires children to have the ability to perceive that there is more than one attributes (physical & behavior, interactions) to given objects. Dixon et al. (2001) identified the following steps that might be followed during the learning of the skill of controlling variables.

i. Have the learners brainstorm to determine the factors that are involved in the investigation.
ii. Ask the learners how they might determine the set-up of the investigation that would result in the maximum solution of the problem. Lead the learners to the conclusion that they will need to compare only one factor at a time.
iii. Before beginning the data collection have learners work in groups to identify the factors that they will keep constant and those that they will

vary during their investigations.

2.4.3 Defining operationally skill

Defining operationally means developing statements that present concrete descriptions of an event by telling someone what to do or what to observe (Chiappetta & Koballa, 2002). Defining operationally is the process by which a scientific term is framed in terms of a student's or researcher's own experiences. An operational definition is also used to define what you will do and what you will observe when conducting an investigation. According to Ziman (2000), operational definitions need to be clear and precise so that a reader knows exactly what to observe or measure. Variables can be defined operationally by applying some kind of measurements (measured operational definition) or by listing steps taken in an experiment to produce research conditions (experimental operational definition) (Ary et al. 1990). Operational definitions are necessary so that other scientists will know exactly what the dependent variable is and how it was measured.

Few studies exist which explains how a teacher could help students define experimental variables operationally. Pratt and Hackett (1998) suggest that, by learning science by inquiry, can facilitate the development of defining operationally variables and acquisition of science process skills in general. Teachers are taught inquiry teaching strategies by engaging in inquiry science activities and extending their understanding of the science concepts that they teach (Hyman & Shephard, 1980). Recognizing the importance of developing defining operational skills of hypothetical constructs to students, Specht (2004) created an in-class exercise he calls “Bucket o’ constructs” study. In the exercise, students work in pairs and each pair grabs a slip of paper from a container and they are supposed to do the following;

i. Look over the written plan for carrying out an investigation, or write up a plan.
ii. Identify and list any variables or terms that do not have a single, clear, obvious meaning.
iii. If there are several reasonable ways to make an observation or to perform an action, choose one that suits the purpose of the investigation.

iv. Write a clear, complete definition of what the researcher should do or measure.
v. Check your definition by asking yourself, Will this definition tell another person what to observe or how to measure? If necessary, revise your definition before starting your investigation.

According to Harlen (2000), teachers can facilitate the development of defining operationally variables and other science process skills in general by; (i) providing a variety of materials and resources to facilitate students' investigations, (ii) posing thoughtful, open-ended and authentic questions, (iii) encouraging dialogue among students and with the teacher, and (iv) keeping students' natural curiosity alive during teaching.

2.4.4 Analyzing and interpreting data skill

Data analysis and interpretation is the process of assigning meaning to the collected information and determining the conclusions, significance, and implications of the findings. According to Berg & Philips (1994), interpreting data involves organizing, analyzing, and synthesizing data using tables, graphs, and diagrams to locate patterns that lead to the construction of inferences, predictions, or hypotheses. Harlen (2000) on the other hand defines data interpretation skill as the ability of putting results together so that patterns or relationships between them can be seen. It involves organizing data and drawing conclusions from it. The purpose of the data analysis and interpretation is to transform the data collected into credible evidence and drawing conclusions from it. The skill also involves creating or using tables, graphs, or diagrams to organize and explain information (Rambuda & Fraser, 2004). When analyzing data (whether from questionnaires, interviews, focus groups, or whatever), always start with a review of your research goals, i.e., the reason you undertook the research in the first place (Berg & Philips, 1994). Leinhardt et al. (1990) argued that students are more motivated to develop their own analysis when they understand that interpretations other than those of the authors may be valid. Substantial studies on the teaching of data analysis and interpretation skill to students exist. According to Harlen (2000), the central part of teacher’s role in developing interpreting skills is to ensure that results are used and students don't rush from one activity to another without talking about and thinking through what the results mean.

To engage in thinking about and discussing data, teachers need to be able to move back and forth between tabular and graphic data representations and verbal statements about the data Similarly, Leinhardt et al. (1990) claimed that students should learn to question the assumptions underlying data collection, analysis and interpretation, and the reasonableness of inferences and conclusions. A substantial body of literature has indicated that even the most routine analysis of data that has information embedded within the graphics may be difficult for older children and even university students (Goldberg & Anderson, 1989) to interpret. Goldberg & Anderson (1989) believe that without a sound understanding of the subtleties (and differences) that underpin different representations, it is difficulty for students to develop sense making within different mathematical contexts.

2.4.5 Experimenting skill (the skill of designing experiments)

An experiment is a test under controlled conditions that is made to demonstrate a known truth, examine the validity of a hypothesis, or determine the efficacy of something previously untried (Ziman, 2000). Experimenting, on the other hand, is an ability of being able to conduct an experiment, including asking an appropriate question, stating a hypothesis, identifying and controlling variables, operationally defining those variables, designing a "fair" experiment, conducting the experiment, and interpreting the results of the experiment (Padilla,1990). An opportunity to practice all individual science process skills discussed above (from section 2.4.1, 2.4.2, 2.4.3 and 2.4.4) is provided by an experiment. Experiments according to Ary et al. (1990) are scientific investigations in which the researcher controls some independent variables and observes the effects of these manipulations on the dependent variable.

Before experimenting, the investigator starts with a question which needs to be solved and then identifying the variables that need to be held constant (Ary et al. 1990; McMillan & Schumacher, 1997; Rezba et al. 1995). According to these authors, the two steps above are followed by formulation of hypotheses to be tested, defining variables operationally, designing an investigation, rerunning trials and finally collecting data for interpretation. Research indicates that activity-based and hands-on activities increase experimenting skill proficiency in processes of science, especially laboratory skills and specific science process skills, such as experimenting graphing and interpreting data (Shymansky & Penick; 1981, Mattheis & Nakayama, 1988; Rutherford, 1993; Rowland, 1990). A study conducted by Shymansky & Penick (1981) concluded that activity-centered classrooms encourage student creativity in problem-solving, experimentation of ideas, promote student independence, and help low ability students overcome initial handicaps. Walters & Soyibo (2001) suggests that students´ experimenting skill is greatly developed when a science teacher guides them in i. stating the hypothesis that is testable. ii. writing out detailed steps to their procedure. iii. determining the independent and the dependent variables. iv. including a description of their control and how it served as a control. v. including a description of their experimental groups.vi identifying factors that must remain constant throughout the experiment.vii designing a data table. viii. graphing their findings and lastly, ix. forming a conclusion based on the data gathered. Zimmerman (2007) on the other hand suggested that where appropriate, students should be included in determining the lab problem, inventing experimental procedures and designing the actual experiment.

2.5 Measurement of science process skills using paper-and-pencil test

2.5.1The history of science process skills measurement using paper-and-pencil test

The science educational reforms of the 1960's and 1970's prompted the need to develop various test instruments for assessing both teachers and students' level of science process skills (Dillashaw & Okey, 1980 & Mungandi, 2005). Since then, scientific process skills evaluations performed through paper and pencil multiple-choice test have been very famous (Burns et al. 1985; Dillashaw & Okey, 1980). Historically, there has been a number of paper-and-pencil test have been developed to assess science process skills for elementary and secondary students (see test of integrated process skills by Dillashaw and Okey,1980; integrated science process skill test TIPS II by Burns et al. 985; process skills test by McKenzie and Padilla, 1986; science process skills test by Onwu and Mozube, 1992; integrated process skill by Mungandi, 2005; and the recent one by Shahali & Halim, 2010). When undergoing scientific validation, these paper and pencil science process tests are considered as powerful tools for measuring one's knowledge of science process skills.

However, as pointed out by Shahali & Halim (2010) science process skills test may either be curriculum specific or non-curriculum specific. Researchers who construct curriculum specific tests believe that scientific inquiry cannot be independent of domain knowledge (Millar & Driver, 1987). Dietz and George (1970) for example constructed multiple-choice questions to test the problem-solving skills of elementary students. This test established the use of written tests as a means to measure problem-solving skills. Relay (1972) on the other hand developed a curriculum specific test of science inquiry skills for grade five students which measured science process skills of identifying and controlling variables, predicting and inferring, and interpreting data. Molitor and George (1976) also developed a test of scientific skills (TSPS), which focused on the inquiry skills of inference and verification for grades four to six learners. The test was considered to be valid, but had a low reliability of 0.66 and therefore, the test was discarded.

On the other hand, the science process skills test may be non-curriculum specific. Nehring et al. (2012) argued that scientific inquiry is assumed to be independent of domain knowledge. Among the earlier researchers to develop and validate a noncurricular specific science process skills test were Dillashaw and Okey in 1980. Dillashaw & Okey (1980) is a comprehensive test of integrated science process skills (TIPS) which included most of the integrated skills such as (i) stating and revising hypotheses, (ii) identifying and controlling variables, (iii) operationally defining of critical terms, (iv) graphing and data interpretation, and (v) designing of experiments. This non-curriculum specific science process skills test was developed for middle and secondary school students. The test had 36 multiple-choice items validated for students at secondary school level (Dillashaw & Okey, 1980). The test was designed to be taken in a single, untimed session. Each of the five integrated science process skills is assessed by six items on the test. All items had four response choices, and they have been stated in a practical problem context. Items were drawn from all science content areas to avoid favoring any particular science background (Dillashaw & Okey, 1980). The Cronbach alpha reliability of TIPSI test was established by Dillashaw & Okey (1980) to be 0.89 by using over 700 secondary school students of grades 7-12. Content validity was established using specific objectives judged by a panel of science educators. The panel found a mean score of 18.99 (s.d. 7.60) for students from general curricula. Readability index was assessed and found to be 9.2 (Dyer et al. 2004; Hamilton & Swortzel, 2007). The test is still very much in use to date and it has been employed in several studies.

Burns et al. (1985) revised the TIPSI and developed TIPS II to measure five components of integrated process skills, which are identifying variables, identifying and stating hypotheses, operationally defining, designing investigation and graphing and interpreting data. As a follow up of TIPSI test discussed above, Burns et al. (1985), developed a similar test which was referred to as the Test of Integrated Science Process Skills II (TIPS II). The test was based on the objectives and format of the original TIPSI and the same number of items which is 36. Its reliability was found to be 0.84. The test is also still very useful to date and employed in several studies such as that by Rowe & Foulds (1996) and the one by Ates in (2004). A recent study by Keil, Haney & Zoffel (2009) on improvements in student achievement and science process skills also employed this test to measure the knowledge level of students in science process skills.

2.5.2 Development and validation of science process skills test in Africa

In Africa, Onwu and Mozube (1992) were the early scholars in Africa to develop and validate science process skills test. They used Nigerian secondary education curriculum to develop and validate a science process skills test for secondary school science students. The test was a 36 multiple choice questions intended to measure students' skills in identifying and controlling variables, defining variables operationally, formulating hypotheses, interpreting data and in designing experiments (Onwu & Mozube, 1992). The test was considered valid in the Nigerian education context and had a reliability of 0.84.

However, a recent test of integrated process skills was developed by Mungandi of the joint center for science mathematics and technology education at the University of Pretoria, South Africa in 2005. The test is a thirty multiple-choice and its reliability was estimated using the split-half method with 1043 learners of grade 9, 10, and 11 to be 0.81. The test has been proven to be gender and race neutral. Its reliability coefficient is well above the lower limit of the acceptable range of values for reliability, and it is within the range of reliability coefficients obtained from similar process skills tests, such as that by Dillashaw and Okey (1980) who obtained a reliability of 0.89 and that by Burns et al. (1985) who also obtained a reliability of 0.84. The readability level of this instrument was found to be 70.29 (Mungandi, 2005). This high readability value of the instrument implies an easy to read and understand test (Zeitler, 1981). The test has an internal validity of 0.97 and a concurrent validity of 0.56.

2.5.3 The debate for using paper and pencil test in assessing process skills

Curricular reforms towards science process skills development entail the need for having evaluation methods which would measure students' level of acquisition of these skills. Multiple-choice tests for assessing students' performance on SPSs have been commonly used. However as Shavelson et al. (1992) and many other researchers maintain, multiple choice format tests are not effective in measuring complex problem-solving skills, science process skills, divergent thinking, and collaborative efforts among students. As a result of the weaknesses of multiple choice format, some researchers have discussed the need for having a more authentic assessment of students' process skills (Souchek & Meier, 1997; Rebza et al. 1995; Harlen, 1999). Souchek & Meier (1997) for example have argued that in order to determine the change of the students' SPS, one should asses extend to which the students have acquired and use the science process skills in the novel learning situation (Souchek & Meier, 1997). According to Harlen (1999), the best way to measure the science process skills of students is laboratory reports, oral presentations, and observation. In authentic assessment, students are asked to perform certain hands-on activities developed by the researcher to assess SPSs through experimentation (Ayala et, al. 2002; Klein et al. 1998; Rezba et al. 1995).

As it has been described above, researchers who criticize paper and pencil multiple choice test in measuring students SPS suggests either authentic or observation methods. Authentic testing is a form of assessment in which students are asked to perform real-world tasks that demonstrate meaningful application of essential knowledge and skills. Stinggins (1987) described this method as assessments that call upon the examinee to demonstrate specific skills and competencies, that is, to apply the skills and knowledge they have mastered. Omoifo and Oloruntegbe (2000) also suggested that on the spot assessment is an additive to paper and pencil test technique for assessing science process skills. Harlen and Jelly (1997) developed observation criteria for each skill in order to determine the improvement of students' science process skills. As the name suggested in using observation method, the assessors observe the students performing the task and see if they have the ability to perform it properly. It is the most obvious form of assessment where an educator watches someone doing something to see if they can do it properly. According to Harlen and Jelly (1997), observation is most effective when it follows a systematic plan. This might involve, for instance, seeing and recording which students use physical materials, which do most of the problems mentally, which use thinking strategies, and which rely on memorized facts (Harlen & Jelly, 1997).

However, the authentic methods for measuring sapience process skills such as through laboratory practical work and physical observation described above have a number of constraints particularly in the context of teaching and learning in large under-resourced science classes (Onwu & Stoffels, 2005; Mungandi, 2005). In this context, a valid, reliable and discriminatory paper-and-pencil test becomes the only option. When scientifically validated, paper and pencil science process tests are considered as powerful tools for measuring one's knowledge of science process skills. For example, Sanchez & Betkouski (1986) used multiple regression analysis to examine 16 predictors of students' success in Chemistry classes and found that scores on the test of science process skills ranked third in predicting students' final grades in Chemistry, below GPA and sex. They are easy to score, and when constructed well, assume many of the psychometric properties that characterize valid assessment practices.

2.6 Criteria and procedures in the development and validation of a science process skills test

When constructing a good test, educators and researchers are guided by a number of principle concerns. In general, they wish to maximize discrimination, score variance, the degree of reliability, and evidence to support validity claims. A test's usefulness, according to Bachman (2000), can be determined by considering the following measurements qualities of the test: reliability, validity, discrimination power, authenticity, and practicality. The general trend in the test development and validation has been (i)Deciding on the principles upon which the constructed test will be based(ii) defining constructs and objectives to be measured (iii) collecting and preparation of test items (iv)conducting initial validation (v) carrying out pilot testing (vi)performing item analysis, (vii) establishing reliability criteria and (viii) conducting validation (Spector, 1992; Bachman, 2000; Onwu & Mozube, 1992; Burns et al. 1985; Dillashaw & Okey, 1980). According to Mungandi (2005), a valid and reliable test should have characteristics that fall within the acceptable range of values for each characteristic such as validity, reliability, discrimination index, index of difficulty and readability. This section discusses literature and studies on these test quality criteria.

2.6.1 Test validity criterion

Validity is arguably the most important criteria for the quality of a test. Lynn (1986) defined validity as refers to whether or not the test measures what it claims to measure. It refers to the degree in which the test or another measuring device is truly measuring what we intended it to measure. For some time, test validity has been broken into i. content validity, ii. concurrent criterion-related validity (predictive and concurrent) ,iii. construct validity and iv. face validity (Shultz & Whitney, 2005; Haynes et al. 1995; Lynn, 1986; Waltz et al. 2005). Content validity is the extent to which the elements within a measurement procedure are relevant and representative of the construct that they will be used to measure (Haynes et al. 1995). According to Reckase (1998), content validity is a non-statistical and the most important type of validity as it examines test content to determine whether it covers a representative sample of the behavior domain to be measured. Generally, establishing content validity is a necessarily initial task in the construction of a new measurement procedure (Waltz et al. 2005). However, as noted by Lynn (1986), this kind of validity is often measured by relying on the knowledge of people, the experts who are familiar with the construct being measured. Establishing strong support for content validity is a challenge because many reviewers are needed to avoid an inflated estimate of validity that often results when experts endorse most items.

Criterion-related validity, on the other hand, pertains to evidence of a relationship between the attributes in a measurement tool with its performance on some other variable (Waltz et al. 2005). It is indicated when measures on the predictor and the criterion variables are correlated and the strength of the correlation substantially supports the extent to which the instrument estimates performance on each criterion (Waltz et al. 2005). Two types of criterion validity are predictive and concurrent validity (Shultz & Whitney, 2005. According to Shultz & Whitney (2005), predictive validity concerns on how well an individual's performance on an assessment measures how successful he will be on some future measure. It is the degree to which test scores predict performance on some future criterion (Shultz & Whitney, 2005). High correlations between the original measure and criterion variables reinforce the conclusion that the tool is a valid predictor of the specified criteria. Concurrent validity refers to how the test compares with similar instruments that measure the same criterion (Waltz et al. 2005).

Another important type of validity is the so-called construct validity and several studies with regard to the study of this validity have been published (Messick, 1980; Ross, 1998; Reckase, 1998; Yen, 1998). Construct validity is considered an overarching term to assess the measurement procedure used to measure a given construct because it incorporates a number of other forms of validity (i.e., content validity, convergent and divergent validity, and criterion validity) that help in the assessment of such construct validity (Messick, 1980). Messick, (1980) defines construct validity as the experimental demonstration that a test is measuring the construct it claims to be measuring. Such an experiment according to Messick (1980) could take the form of a differential-groups study, wherein the performances on the test are compared to two groups: one that has the construct and one that does not have the construct. If the group with the construct performs better than the group without the construct, that result is said to provide evidence of the construct validity of the test (Reckase, 1998). Convergent validity and discriminant validity are commonly regarded as subsets of construct validity. John & Benet-Martinez (2000) describe convergent validity tests that constructs that are expected to be related are, in fact, related while discriminant validity (or divergent validity) tests that constructs that should have no relationship do, in fact, not have any relationship. Face validity is the last and the weakest form of validity (Nevo, 1985). Schultz & Whitney (2005) defines Face validity as the extent to which a test is subjectively viewed as covering the concept it purports to measure. In other words, a test can be said to have face validity if it "looks like" it is going to measure what it is supposed to measure by observant. Face validity could easily be called surface validity or appearance validity since it is merely a subjective, superficial assessment of whether the measurement procedure you use in a study appears to be a valid measure of a given variable or construct (Nevo, 1985).

In estimating test validity, validity coefficient is calculated as a correlation between the two items being compared (Kaplan & Saccuzzo, 2001).The validity coefficient is a statistical index used to report evidence of validity for intended interpretations of test scores and defined as the magnitude of the correlation between test scores and a criterion variable (i.e., a measure representing a theoretical component of the intended meaning of the test) (Moss, 1998 & Reckase, 1998). Different researchers have different views on the acceptable range of test validity coefficient. For example Kaplan, & Saccuzzo (2001) a validity coefficient of 0.6 and above is considered high, which suggests that very few tests give strong indications of validity evidence. Gall & Borg (1996) and Hinkle et al. (2003) recommends a validity coefficient of 0.7 and above as a suitable a suitable value for standard tests. Adkins (1974) on the other hand stated that appropriateness of validity coefficient depends on several factors and that coefficient of a unit or close to a unit.

2.6.2 Test reliability

Reliability is another important component of a good psychological test and it looks consistency or reproducibility of test scores. There are numerous ways to define and estimate test reliability (Kaplan, & Saccuzzo, 2001; Rudner & Schafer, 2001; Messick, 1980; Moskal & Leydens, 2000). Rudner & Schafer (2001) defined reliability as the degree to which test scores for a group of test takers are consistent over repeated applications of a measurement procedure and hence are inferred to be dependable and repeatable for an individual test taker. Rudner and Schafer (2001) argue that the best way to view reliability is the extent to which test measurements are the result of properties of those individuals being measured. More important to understand is that reliability estimates are a function of the test scores yielded from an instrument, not the test itself (Moskal & Leydens, 2006; Henson, 2001). Accordingly, reliability estimates should be considered based upon the various sources of measurement error that will be involved in test administration (Crocker & Algina, 1986; Henson, 2001).

There are many methods to estimating the reliability of a measurement tool, each resulting in a different dimension of reliability. According to Crocker & Algina (1986), these methods include test-retest method, alternative (parallel form) reliability, inter-rater reliability and internal consistence reliability. The application of these methods will vary depending on your testing situation and how you plan to use the test results. Moskal& Leydens (2006) describe test-retest reliability as a measure of reliability obtained by administering the same test twice over a period of time to a group of individuals. To gauge test-retest reliability, the test is administered twice at two different points in time. This kind of reliability is used to assess the consistency of a test across time. The first round test should be administered, a sufficient period of time should elapse, and the test should then be administered once again (Crocker & Algina, 1986; Gregory, 1992).

Alternate form reliability, on the other hand, is a measure of reliability obtained by administering different versions of an assessment tool (both versions must contain items that probe the same construct, skill, knowledge base, etc.) to the same group of individuals (Moskal & Leydens, 2006). According to DeVellis (2011), the scores from the two versions can then be correlated in order to evaluate the consistency of results across alternate versions .This approach is particularly useful in the context of standardized testing procedures, where it is ideal to have multiple, and equivalent, forms of the same test. In this technique, a coefficient of equivalence is yielded (Crocker & Algina, 1986; DeVellis, 2011; Gregory, 1992).

Another important way of estimating reliability is through Inter-rater reliability. Crocker & Algina (1986) defined inter-rater reliability as a measure of reliability used to assess the degree to which different judges or raters agree in their assessment decisions. Gregory (1992) argued that inter-rater reliability is useful because human observers will not necessarily interpret answers the same way; raters may disagree as to how well certain responses or material demonstrate knowledge of the construct or skill being assessed. The scores from the two raters can then be correlated in order to evaluate the consistency of results (DeVellis, 2011).

The last type of reliable estimate is known as internal consistency reliability. Crocker & Algina (1986) describe internal consistency as a form of reliability which is used to judge the consistency of results across items on the same test. Popham (2002) reminds us that this type of reliability necessitates only one test administration. High reliability estimates will result into high inter-item correlations among the items or subscales (Crocker & Algina, 1986; DeVellis, 2011; Gregory, 1992). Although there are three different measures of estimating internal consistency reliability of a research tool, the most widely used measure is Cronbach's coefficient alpha. Cronbach's alpha is actually an average of all the possible split-half reliability estimates of an instrument (Crocker & Algina, 1986; DeVellis, 2011; Gregory, 1992; Thompson, 1994). The two lesser-used techniques of estimating coefficient alpha are appropriate in limited circumstances. For example, the Kuder-Richardson 20 is appropriate for use with dichotomously-scored items, and Hoyt's method is useful in particular testing situations that involve computer programming (Crocker & Algina, 1986).

Another type of internal consistency measure is split-half reliability. This technique literally takes an instrument, assesses the reliability of the first half, and then compares this estimate to the reliability measure of the second half (Krzanowski & Woods, 1984). It should be noted that reliability estimates are often underestimated when computing split-half reliability, due to the shortened nature of the instrument (Krzanowski & Woods, 1984). This error in calculation can be addressed by using the Spearman-Brown prophecy, which provides the means necessary to estimate reliability for the full-length test based on split-half calculations (Crocker & Algina, 1986; Thompson, 1994).

As it is for validity, different researchers also hold different views in regard to the acceptable ranges of reliability. According to Nunnally & Bernstein (1994), the acceptable reliability estimates of social sciences research instruments ranges from 0.70 to 0.80. However, research in the physical sciences typically demands more rigorous reliability standards, as the constructs involved are more concrete and easily defined. In both settings, acceptable reliability estimates should be congruent with the implications of the test scores. That is, higher stakes testing should have higher standards of instrument reliability (Nunnally & Bernstein, 1994).

2.6.3 Standard error of measurement (SEM)

Reliability of a measuring instrument can also be expressed in terms of a statistic know as standard error measurement (Gay & Airasian, 2000). The standard error of measurement (SEM) is defined in the standards for educational and psychological testing (1985) as the standard deviation of errors of measurement that is associated with the test scores for a specified group of test takers. It is one of the common test validation statistical criteria. According to Gay & Airasian (2000), SEM measures the variability of the errors of measurement and is directly related to the error score variance. It estimates how repeated measures of a person on the same instrument tend to be distributed around his or her "true" score. Mungandi (2005) contends that standard error of measurement helps us to understand that the scores the scores obtained on the educational measure are only estimates, and may be considerably different from an individual´s presumed true score. Since all measurement contains some error, it is highly unlikely that any test will yield the same scores for a given person each time they are retested. The true score is always an unknown because no measure can be constructed that provides a perfect reflection of the true score. SEM is directly related to the reliability of a test; that is, the larger the SEM, the lower the reliability of the test and the less precision there is in the measures taken and scores obtained (Gay & Airasian, 2000). In this study, Standard error of measurement will be determined to further estimate the reliability of science process skills test developed.

2.6.4 Readability of the test

Readability is the level of ease or difficulty with which text material can be understood by a particular reader who is reading that text for a specific purpose. It is one of the commonly cited as criteria for a good test. According to Richards et al. (1992), readability means how easily written materials can be read and understood. Bailin & Grafstein (2001) asserts that communication presupposes comprehension, but the increasing variety, volume, and complexity of written materials make understanding more and more of a problem. Hence the main purpose of readability according to Hewitt & Homan (2004) is to ensure that a given piece of writing reaches and affects its audience in the way that the author intends. On that note, readability studies concentrate on the linguistic factors, in particular, word length and sentence length.

There is a general agreement among linguists that readability depends on several factors such as the average length of sentences, the number of new words contained, and the grammatical complexity of the language used in a passage (Bailin & Grafstein, 2001; Hewitt & Homan, 2004; Wright, 1982). However, Harris & Hodges (1995) adds that readability is dependent upon not only characteristics of a text but also characteristics of a reader. Thus, one important characteristic of a useful, informed definition of readability is that it reflects the interactive nature of the construct. Text and reader variables interact in determining the readability of any piece of material for any individual reader. Chall & Dale (1995) concludes that the purpose of readability assessment is to effect a 'best match' between intended readers and texts . . . thus, the optimal difficulty comes from an interaction among the text, the reader, and his/her purpose for reading".

Readability tests, readability formulas, or readability metrics are formulae for evaluating the readability of text, usually by counting syllables, words, and sentences. Authors and publishers utilize readability indices to quantify the reading grade level of which a typical student can read a text. Over the past 8 decades, more than 200 readability indices have been proposed and utilized in various literary contexts (DuBay, 2004). The commonly used however include Fry Graph Readability Formula (1969), SMOG readability formula (1969), Flesch –Kincaid Reading Ease Readability Formula (1948), the Dale-Chall (1955), Powers-Summer-Kearl (PSK), Spache (1953) etc .This study is going to adopt the popular readability formula by Flesch-Kincaid (Flesch, 1948) which uses factors such as the number of sentences in a passage and the syllable count of the words in the passage. In readability assessment, the higher the readability score, the easier the text is to understand and the vice versa is true. The recommended range of scores for a test instrument is 60-70, which is the plain English level (Mungandi, 2005).

2.6.5 Test item analysis

Another consideration for a quality test is an analysis of its items. The quality of an item decides the quality of the overall test. Items within a measure differ in terms of their difficulty and discrimination abilities. Classical item analysis helps in improving the quality of tests by revising and improving the items in the test (Johnson et al. 2000, Sherry, 1997; Allen & Yen, 2001; Nunnally & Bernstein, 1994). Item analysis according to Nunnally & Bernstein (1994) is a process which examines student responses to individual test items in order to assess the quality of those items and of the test as a whole. Item analysis is especially valuable in improving items which will be used again in later tests, but it can also be used to eliminate ambiguous or misleading items in a single test administration. Allen & Yen (2001) argued that item analysis for a multiple choice test always includes three statistics i. difficulty index, ii. discrimination index and iii. distracters analysis.

Difficulty index, or commonly noted as the p-value for dichotomously scored items, is a proportion of the examinees that answered an item correctly (Singamaneni, 2011). This index according to Allen & Yen (2001) represents the level of difficulty based on the particular group of examinees to which the test was administered. In a test, it is important to know whether the difficulty of an item is suited to the level of students for whom the test is intended. Singamanenim (2011) claimed that the p-value is sample dependent and varies when groups of different ability levels are administered the same examination. Difficulty index value ranges from 0.0 to 1.0 and the higher the value, the easier the item is (Singamaneni, 2011). In traditional achievement tests according to Singamaneni (2011), items displaying values closer to 0 (indicating that almost all students got the item wrong) and 1 (indicating that almost everyone got the item correct) should be revised or removed because they offer the little ability to discriminate among students at varying proficiency levels.

As for validity and reliability values, different researchers hold different views on the acceptable range of difficulty index values. For example Singamaneni (2011) stated that items having difficulty ranges from 0.2 to 0.8 provide the maximum information about proficiency among students. The p-value from 0.2 to 0.8 has also been suggested by Nitko (1996) who argued that item having difficulty indices between 0.2 - 0.8 should be retained and all other items out of this range will either be discarded or modified. However, Popham (2002) determined levels and distribution of difficulty recommending that an ideal multiple-choice item for testing should be around 62.5% (p-value = 0.625).

Item discrimination is the other statistic in classical item analysis. Allen & Yen (2001) defined item discrimination as an index of how effectively the item separates examinees who vary in their degree of knowledge tested and their ability to use it. The discrimination index (D) is one of the most useful methods for dichotomously scored items, due to its computational simplicity. Singamanenim (2011) argued that item discrimination index indicates whether items are discriminating students based on their ability to perform That is, the item is able to distinguish between high and low performing students. The value according to Mungandi (2005), is calculated by computing the difference of difficulty indices of higher and lower achievers in each item. Singamanenim (2011) suggested that the value of item discrimination can be stable, using the upper 27% and the lower 27% of examinees, if no distinction is made among the members of each group separately. Zero discrimination occurs when equal numbers in both groups answer correctly. Negative discrimination occurs when more students in the lower group answer correctly than the upper group.

Allen & Yen (2001) suggested a scale for interpreting item discrimination in which items with negative values are judged unacceptable (and should be checked for errors) and those with discrimination values between 0% and 24% are potential candidates for approval. Items in a test having the discrimination index of 0.3 and above have been recommended by several researchers (Adkins, 1974; Hinkle et al. 2003) be good enough to be to be retained. Ndalichako & Rogers (1997) suggests that values below 0.2 are weak, and values above 0.4 are desirable. However, Singamaneni (2011) claimed that items with discrimination values from 25% to 39% are considered good items and those with values at or above 40% are judged to be excellent items.

In general, researchers agree that for an item to be considered appropriate it should exhibit proper difficulty level and discrimination value in terms of the intended purpose of the test (Singamaneni, 2011; Adkins, 1974; Mungandi, 2005; Hinkle, 1998 & Allen and Yen, 2001). However, in the reverse case such as when items have inappropriate difficulty or discriminative values, the items can be improved through the distracter analysis (Singamaneni, 2011). Hills (1981) defined distracter analysis as the process for evaluating whether alternative responses to each item effectively function. Considering the ideal responses to the alternatives of an item, Hills (1981) suggested four patterns: (1) at least one examinee should select every distracter, (2) the right answer should be selected much more frequently by the examinees in the upper group than those in the lower group, (3) each distractor must be chosen more by the lower-scoring examinees than the higher-scoring ones, and (4) it is desirable that the difficulty level of each item is similar to the optimal proportions.

2.7 Review of literature and studies related to inquiry based approach to science

2.7.1 Meaning and nature of inquiry-based approach

The meaning and nature of inquiry-based learning are highly contested among researchers and educationists. The phrase has itself allows many permutations and alternatives. Thousands of of descriptions of what exactly entails and constitutes inquiry-based learning exist throughout educational literature. Multiple interpretations of this multi-faceted approach according to Colburn (2000) have resulted in an overall confusion about the meaning of inquiry and what inquiry implies for the teacher. Terms such as problem-solving approaches, project-based learning, research-based teaching, discovery learning, and inductive teaching have been terms used interchangeably to mean inquiry-based approach (Poon et al. 2004; Spronken-Smith et al. 2007; Anderson, 2002; & Warner and Anna, 2008)

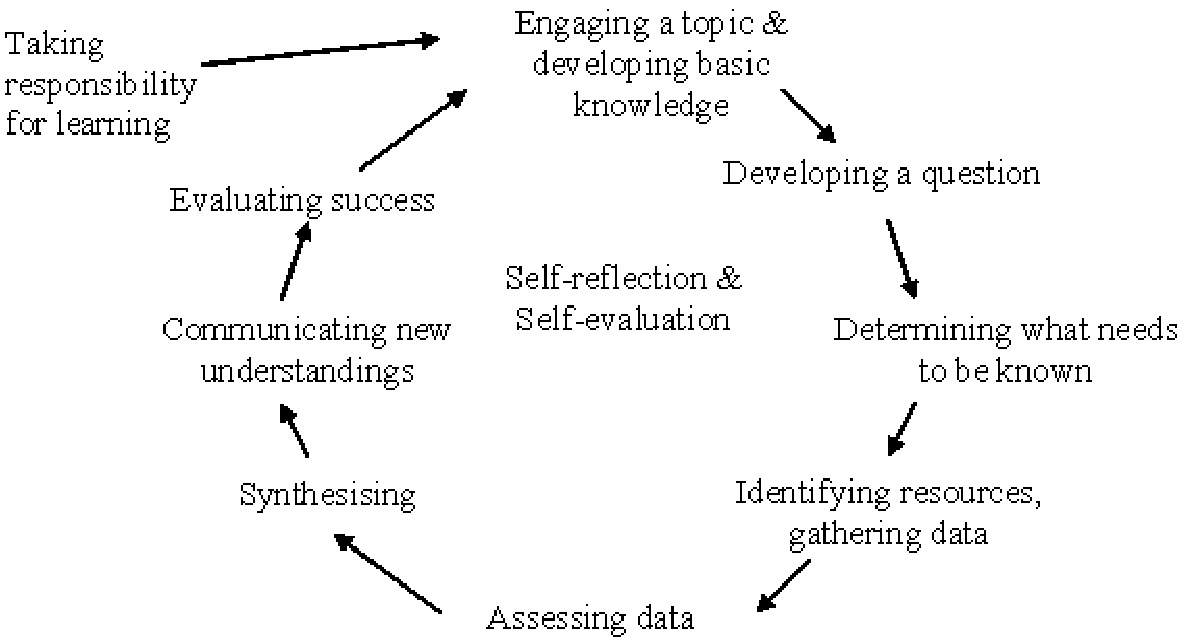
However, with deep analysis of literature, one might find a commonality of opinion among researchers about what constitutes inquiry-based learning (IBL). Spronken-Smith et al. (2007) drew on this commonality to define IBL as a constructivist pedagogy which best enables students to experience the processes of knowledge creation. The approach involves asking questions, gathering and analyzing information, generating solutions, making decisions, justifying conclusions and taking action (Spronken-Smith et al. 2007). The list of core ingredients of an IBL approach that most researchers are in agreement with includes the following,

i. learning stimulated by inquiry, i.e. driven by questions or problems
ii. learning based on a process of seeking knowledge and new understanding
iii. a learning-centred approach to teaching in which the role of the teacher is to act as a facilitator
iv. a move to self-directed learning with students taking increasing responsibility for their learning and the development of skills in self-reflection
v. an active approach to learning (Spronken-Smith et al. 2007).

Historically, inquiry-based learning as an instructional method was developed during the discovery learning movement of the 1960s. According to Edelson et al. (1999), the approach was developed in response to a perceived failure of more traditional forms of instruction, where students were required simply to memorize fact laden instructional materials. Inquiry learning is a form of active learning, where progress is assessed by how well students develop experimental and analytical skills rather than how much knowledge they possess (Edelson et al. 1999). The approach draws upon a constructivist learning theory where understanding is built through the active development of conceptual mental frameworks by the learner.

Another key issue that most of the researchers and their models also agreed is that inquiry-based learning (IBL) is cyclic implying that ending of one inquiry leads to new interests and more questions which lead to another inquiry (Spronken-Bishop et al. 2004; Justice et al. 2007; Hancock et al. 1992 & Bybee et al. 2006). Bishop et al. (2004) for example described inquiry-based learning as a cycle or a spiral, which implies formulation of a question, investigation, creation of a solution or an appropriate response, discussion and reflection in connection with results. A group of McMaster university teachers (Justice et al. 2007) also involved in IBL, developed a circular model of the inquiry process (Figure 2.1). According to the model, students become engaged with a topic, develop a question to explore, determine what information is needed, gather data, synthesize findings, communicate findings and then evaluate the success. Core to the process is an attitude of self-reflection and evaluation, which are seen as both a product of the inquiry and an enabler of success at every stage (Justice et al. 2007).

Figure 2.1 below: Model of the inquiry process (Justice et al. 2007)



2.7.2 Advantages of using inquiry based learning

There thousands of evidences from researchers all over the world in support for the use of Inquiry-based approach from studies which have evaluated the impact of this mode of teaching on student learning outcomes. Prince & Felder (2006) for example provide a good overview of four studies which evaluated the effectiveness of inquiry based approaches to students (Haury, 1993; Rubin & Norman, 1992; Shymansky, 1990; Crawford, 2000; Prince & Felder, 2006). Prince & Felder (2006) metaanalytic study concludes that IBL is generally more effective than traditional teaching for achieving a variety of student learning outcomes such as academic achievement, process skills, analytic abilities, critical thinking and creativity. In his study on the effectiveness of inquiry based instruction, Germann and Aram (1996) found that the directed inquiry approach is effective in learning science process skills and scientific problem – solving.

The purpose of Gregg's study (1995) was to identify interaction patterns that emerged during mathematics instruction in elementary school classrooms that established an inquiry mathematics tradition. Preliminary analysis from this study suggested that aspects of an inquiry approach to Mathematics instruction had a positive impact in providing gender-equitable learning opportunities for boys and girls. In his study, Crawford (2000) discusses how the IBL approach results in students acting as apprentice researchers in the field. The research result supported the use of inquiry based techniques to enhance students´ critical thinking. Others studies (Carin & Bass, 2001; Kyle, Bonnstetter, & Gadsden, 1988; Brew, 2003; Gregg, 1995) shows that inquiry-based instruction develops reasoning skills and heightens students' motivation toward science.

Some studies compared the learning outcomes of students taught with traditional approaches with those taking an IBL version of the same course (Justice et al. 2007; Berg et al. 2003; Wolf, 1993). Justice et al. (2007) for example used five years of data to examine whether taking a first year IBL course made a difference in students' learning and performance. In a comparative study between students taking an IBL course and those who did not, and, taking into consideration factors such as age, gender, high-school grade point averages etc., they found that students who took the inquiry course had statistically significant positive gains in passing grades, achieving Honours and remaining in the university. Berg et al. (2003) on the other hand compared the learning outcomes of an open-inquiry and an expository version of a first year chemistry laboratory experiment. Data on student experiences of the two approaches were gathered from interviews, questions during the experiment and students' self-evaluations. The key findings of this study were that students taking the open-inquiry experiment version had more positive outcomes including a deeper understanding, higher degree of reflection, the achievement of higher order learning and more motivation. Change & Mao (1998) investigated the effects of an inquiry-based instructional method on earth science students' achievement. Their result concluded that the inquiry approach provided more opportunities for students to apply intellectual skills than expository instruction.

Other studies opted to compare how the inquiry based learning experience has changed students' perceptions, motivation and interests about the topic (Kennedy & Navey-Davis, 2004; Shymansky et al. 1990; Haury, 1993; Houlden et al. 2004; Change & Mao, 1998). In its essence, Haury (1993) found that inquiry-oriented teaching engages students in investigations to satisfy curiosities, with curiosities being satisfied when individuals have constructed mental frameworks that adequately explain their experiences. Houlden et al. (2004) on the other hand examined medical students' perceptions of an undergraduate research elective and its impact on their learning. They found that the IBL elective resulted in students being more confident in their ability to pursue a research career as well as being more interested in such an option. The study by Marshall & Horton (2011) found that students taught by inquiry approach develop mutual enjoyment of the approach, even if there may be some adjustment and initial anxiety about learning or teaching and enthusiastic for more inquiry courses. For example, Shymansky et al. (1990) found an improvement in achievement, attitude and process skills due to inquiry-based teaching. They also found that the shared inquiry between teachers and students is inherently motivating and supporting students' intrinsic motivation (Shymansky et al. 1990).

2.7.3 Negative aspects associated with the use of inquiry- based learning (IBL)

The science education community has published a wide range of findings of inquiry-based teaching and learning including inconclusive and mixed results with respect to its effectiveness (Colburn, 2008). Few studies have also reported barriers and negative aspects associated with IBL (Crawford, 2007; Reiser, 2004; Burris & Garton, 2007). Justice et al. (2003) for example found that students perceived an increased workload in IBL courses, while Lukie et al. (2004) and Plowright and Watkins (2004) argued that IBL causes anxiety on the part of learners over the need to become self-directed. Another issue regarding inquiry-based learning has to do with a misconception about when to do inquiry. According to Justice et al. (2003), an inquiry is not only done in the laboratory or group work as it is perceived by many teachers it can also be done in lectures that provoke students to think and question.

Several researchers who compared the effectiveness of inquiry approach against conventional approaches have failed to identify the tangible value of IBL as claimed by its proponent. Many of those who critique inquiry-based learning do so because it requires less direct interference from the teacher in the child's learning and its consumption of time both on the part of the teacher and student also. For example in their comparison study, Burris & Garton (2007) found that students taught by traditionalist methods tended to score higher on content knowledge assessments than students taught by constructivist methods. The students taught by the traditionalist method gained an average of nine (9) points from pretest to posttest, whereas, the students taught by the constructivist methods showed an improvement of just over 4 points of their pretest scores. Similar conclusions have been reported by Chall (2000) and Furtak (2006). Kirschner et al. (2006) reviewed a small number of studies that they argue provide evidence against the effectiveness of inquiry-based materials and teaching. The studies they reviewed include some that showed how pure inquiry discovery teaching methods can lead to frustration, some that showed how discovery learning is inefficient because it can lead to false starts and some that found support for direct instruction over discovery learning. Moreno (2004) concludes that students learn more deeply from strongly guided learning than from discovery and inquiry mode. Kirshner et al. (2006) reached a conclusion that minimally-guided instruction is less effective and less efficient than instructional approaches that place a strong emphasis on guidance of the student learning process.

Other educators have criticized the modality used in implementing inquiry-based learning approach. For example researches (Kirshner et al. 2006; Moreno, 2004) have argued that the modality of inquiry-based teaching is very time-consuming. Kirshner et al. (2006) on the other hand has argued that inquiry-based instructions take a lot more effort for a teacher to scaffold a lesson than to simply give students the required information. The authors further argue that inquiry-based learning focuses more on the method of learning and less on the specific content to be learned (Kirshner et al. 2006). Plowright & Watkins (2004) also noted that with an inquiry-based approach, the student always faces difficulties in coping with group dynamics. Thus, despite Anderson's (2002) claim that research has generally shown inquiry teaching to produce positive results, challenges that demand attention still appears to exist. For example in a replication study by Sweller (1994) on learning difficulty and instructional design found that students learn to become better at solving mathematics problems when they study worked-out examples rather than when they solely engage in hands-on problem solving through inquiry. Burris & Garton (2007) concluded that although there is agreement on the contribution of constructivist approaches to factors such as knowledge retention, student satisfaction, motivation, and critical thinking, there is much less agreement on its role in knowledge acquisition (Burris & Garton, 2007).The current researcher holds the view that teaching with the inquiry-based approach is more effective when the students already have strong knowledge of the subject matter at hand.

2.7.4 Teacher and student roles in inquiry based approach

Education literature stated numerous roles that teachers should engage when using the inquiry model of teaching (Crawford, 2007; Crawford 2000; Brew, 2003; Scott 1994 & Anderson, 2002). Crawford (2000) for example identified teacher roles in the inquiry-based lesson as a motivator, diagnostician, guider, innovator, experimenter, researcher, modeler, mentor, collaborator and a learner. All these roles according to all (Crawford 2000) correspond with the principles of constructivist teaching where students are supposed to take lead and ownership of their own learning. In inquiry-based learning according to tea Brew (2003) teachers are no longer function as the all-knowing authorities, imparting knowledge to the unknowing rather, teachers and students are viewed as interactants in posing questions and seeking answers. The teacher thereby assumes the role of facilitator observing and guiding students as the latter engage in processes of knowledge discovery (Crawford 2000; Brew, 2003).

As a result of changes in teacher roles in inquiry, there has been a shift from a traditional, predominantly teacher-centered classroom to a more student-centered classroom, where learners take an active role in their own education. Educators who make this shift no longer spend most of their time in front of the classroom lecturing to students, expecting them to absorb the prepared information. Instead, teachers circulate among the students, listening to them and guiding them with carefully crafted questions, modeling the behavior of scientists as the students are encouraged to engage in authentic scientific research (Crawford 2007). The teacher scaffolds learning for students, gradually removing the scaffolding as students develop their skills.

As for the role of teachers, the role of students also changes drastically when they are learning through an inquiry-based approach. Inquiry learning is a process in which students actively engage in within the science classroom (Anderson 2002). According to Reiser (2004) in IBL students are expected to assume the role of little scientist as they develop their curiosity and inventiveness in posing questions and seeking answers about the natural world. According to Reiser (2004) inquiry based learning requires students work cooperatively in groups, forming hypotheses, designing experiments, observing and analyzing results, classifying, drawing conclusions, and communicating all of the above, as well as critiquing each other's work (Crawford 2000; Reiser, 2004). As the students acquire the skills and concepts of the scientific world, they also become more autonomous learners in both thought and action (Reiser, 2004; Kirshner et al. 2006; Furtak, 2006. The active student engagement in inquiry-based science education described above has a corresponding relationship to teachers' roles.

2.8 Review of literature related to students motivation in learning science

2.8.1 The concept and types of motivation

Educational psychologists have for a long time recognized the importance of motivation in enhancing and supporting student learning. Wigfield et al. (2004) stated that motivation is of particular interest to educational psychologists because of the crucial role it plays in student learning. Several researchers have come up with definitions, categories, orientations and theories of motivation (Dörnyei & Otto, 1998; Ryan and Deci 2000; Seligman & Csikszentmihalyi, 2000; Goldberg & Cornell, 1998; Weinburgh & Englehard,1994; Wigfield et al. 2004). According to Dörnyei & Otto (1998) motivation is the dynamically changing cumulative arousal in a person that initiates, directs, coordinates, amplifies, terminates, and evaluates the cognitive and motor processes whereby initial wishes and desires are selected, prioritized, operationalized and (successfully or unsuccessfully) acted out. Seligman & Csikszentmihalyi (2000) on the other hand described motivation as the driving forces which and helps causes people to achieve goals. It involves a constellation of beliefs, perceptions, values, interests, and actions that are all closely related (Gottfried, 1990). As a result, various approaches to motivation can focus on cognitive behaviors (such as monitoring and strategy use), non-cognitive aspects (such as perceptions, beliefs, and attitudes), or both. In addition, Gottfried (1990) defines academic motivation as the enjoyment of school learning characterized by a mastery orientation; curiosity; persistence; task-endogeny; and the learning of challenging, difficult, and novel tasks. In the social-cognitive theory of human learning (Bandura, 1986), students' characteristics, behaviors, and learning environments are viewed interactively. Motivation is considered as a critical determinant of students' classroom learning and achievement partly because students who are more highly motivated tend to provide greater effort and persist longer at academic tasks than do students who are less motivated (Wolters & Rosenthal, 2000). In cognitive models of motivation, this greater effort and persistence for academic tasks is thought to result mainly from various beliefs, attitudes, and perceptions of the student (Weinburgh & Englehard, 1994).

2.8.2 Importance of motivation in science learning

Generally all researchers agreed that motivation is an important factor in student learning; it is positively correlated with students' willingness to learn, high-level cognition, creativity and performance (Benabou & Tirole, 2003; Broussard & Garrsion, 2004; Johnson, 1996; Lavigne et al. 2007; Tuan et al. 2005; Glynn & Koballa, 2006). Motivated students achieve academically by strategically engaging in behaviors such as class attendance, class participation, question asking, advice seeking, studying, and participating in study groups (Pajares, 2001; Pajares, 2003). According to Tuan et al. (2005) this greater effort and persistence for academic tasks is thought to result mainly from various beliefs, attitudes, and perceptions of the student. The motivated student has the inner strength to learn, to discover and capitalize on capabilities, to improve academic performance and to adapt to the demands of the school context (Glynn & Koballa, 2006; Tuan et al. 2005).

Motivation and its importance for science learning have been widely discussed by science education researchers (Tuan et al. 2005; Wolters & Rosenthal, 2000;

Weinburgh & Englehard, 1994). Many studies have investigated the relations between students' motivation to learn and their achievement in science, mainly their performance and scores in science tests. Kremer and Walberg (1981) for example reviewed 20 studies dealing with student motivation and concluded that there is a positive relationship between motivational constructs and science learning. Later et al. (1985) in a study that analyses various affective determinants, found that the highest correlate to achievement in science was student motivation. Evidence suggest that decisions to engage in effortful learning in science may be influenced by individual students' motivation, including their goals for engaging in an activity, their beliefs about their abilities and the nature of the task, and their valuing of the task (Broussard & Garrsion, 2004). According to Benabou & Tirole (2003) motivation to learn science promotes student construction of their conceptual understanding of science. Gottfried (1990) also found a relationship between motivation and achievement, but she maintains that the causal relationship works in the opposite direction. If motivated, students tend to approach challenging tasks eagerly, persist in difficulty, and take pleasure in their achievement (Pajeres, 2003). Osborne et al. (2003) reviewed literature related to students' attitude and academic achievement and concludes that motivation scores of students were correlated to their scores on science attitude and achievement.

2.8.3 Orientations or constructs of students' motivation

The motivational components or orientations that influence students learning and achievement were reviewed recently by Glynn and Koballa (2006), Eccles and Wigfield (2002), Schunk et al. (2008) and Glynn et al. (2011). According to Glynn & Koballa (2006), motivational orientations consist of six basic constructs which include intrinsic motivation, extrinsic motivation, personal relevance, self-efficacy, self-determination, and assessment anxiety. In studying the motivation to learn science according to Glynn & Koballa (2006), researchers examine why students strive to learn science, how intensively they strive, and what beliefs, feelings, and emotions characterize them in this process. These motivational components have also been linked to science learning and they have been studied extensively by education psychologists. This current research determined the influence of inquiry-based approach to the development of students motivational components as described by Glynn et al.(2011).

2.8.3.1 Intrinsic versus extrinsic motivational components

Numerous educational psychologists agree that students' motivation is classified into i. intrinsic and ii. extrinsic motivation (Ryan & Deci, 2000; Pintrich & De Groot, 1990; Osborne & Collins, 2001; Goldberg & Cornell, 1998; Eccles, & Wigfield, 2002; Pintrich & Schunk, 1996). According to Ryan & Deci (2000) intrinsic motivation arises from a desire to learn a topic due to its inherent interests, for self-fulfillment, enjoyment and to achieve a mastery of the subject. Eccles & Wigfield (2002) on the other hand explained that intrinsic motivation is the true drive in human nature, which drives individuals to search for and to face new challenges. Intrinsically motivated students are the ones whose learning goals is mastery of content and skills not as a means to an end but as an end itself (Ryan & Deci, 2000). Their abilities are put to the test and they are eager to learn even when there are no external rewards to be won (Goldberg & Cornell, 1998). This means that intrinsic motivation refers to engagement in an activity with no reason other than the enjoyment and satisfaction of engagement itself.

By comparison, extrinsic motivation according to Ryan & Deci (2000) refers to engagement that provides means to ends that go beyond the engagement itself. It is the motivation to perform and succeed for the sake of accomplishing a specific result or outcome (Pintrich & Schunk, 1996). Education psychology researchers agree that goals of extrinsically motivated engagement might be the attainment of tangible rewards such as money, prizes, or other benefits; intangible rewards such as social approval, a sense of worthiness, or even a sense of conscientiousness; or the avoidance of tangible and intangible punishments such as time-out, scolding, rejection or sense of low self-worth (Benabou & Tirole, 2003; Pintrich & Schunk, 1996; Ryan & Deci, 2000). Extrinsic motivation typically produces immediate results and requires less effort in comparison to intrinsic motivation (Ryan & Deci, 2000). Students who are very grade-oriented are extrinsically motivated, whereas students who seem to truly embrace their work and take a genuine interest in it are intrinsically motivated (Benabou & Tirole, 2003).

Traditionally, educators consider intrinsic motivation to be more desirable and result in better learning outcomes than extrinsic motivation (Pintrich & Schunk, 1996; Benabou & Tirole, 2003; Vallerand & Bissonnette, 1992; Ryan & Deci, 2000). Engagement out of intrinsic motivation requires no external incentives and enhances motivation to engage again in the future. Studies also suggest that engagement out of intrinsic motivation is associated with enhanced comprehension, creativity, cognitive flexibility, achievement, and long-term well-being (Ryan & Deci, 2000). For example, a research study was done by Lens & Rand (1997) concluded that intrinsically motivated students learn independently and always choose to do challenging tasks and integrate their knowledge acquired in school with their experiences gained from outside school. The negative side of extrinsic motivation is the fact that it often distracts students from true independent learning. Another problem with extrinsic motivators is that they typically do not work over the long term. According to Vallerand & Bissonnette (1992) once, the rewards are removed, students lose their motivation. Pintrich & Schunk (1996) on the other hand believe that extrinsic motivational factors can diminish students' intrinsic motivation.

2.8.3.2 Self-efficacy motivational component

Self-efficacy is one of the very important motivational orientations that have been consistently cited by researchers to affect students learning and achievement (see Graham & Harris, 1989; Pajares & Valiante, 1997; Hamilton & Swortzel, 2007; Zimmerman & Bandura, 1994). Bandura (1997) has defined self-efficacy as beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments. Self-efficacy is important because it is the reason behind why we push ourselves to be the best we can be and it motivates us to never give up and to always do our best (Bandura, 1997). These perceptions of self-capabilities or self-efficacy have also been identified as key factors affecting thought patterns and performance in a wide variety of tasks. Graham & Harris (1989) for example believe that self-efficacy perceptions influence choice of activity, task perseverance, the level of effort expended, and ultimately, the degree of success achieved. According to Pintrich & Schunk (1996), people acquire information to appraise self-efficacy from their actual performances, vicarious (observational) experiences, forms of persuasion and psychological symptoms.

During the two decades since Bandura first introduced the construct, the role of self-efficacy has received extensive support from a growing body of findings from diverse fields (see Riggs & Enoch, 1990; Gibson & Demba 1984; Graham & Harris, 1989; Pajares & Valiante, 1997). Self-efficacy has also received increasing attention in educational research, primarily in studies of academic motivation (Pintrich & Schunk, 1996). Researchers have established that self-efficacy beliefs and behavior changes and outcomes are highly correlated and that self-efficacy is an excellent predictor of behavior. When facing difficulties, students who have a high sense of efficacy for learning should expend greater effort and persist longer than those who doubt their capabilities (Ashton & Webb, 1984; Schunk & Swartz, 1993). Percepts of self-efficacy also influence level of skillful performance (Schunk & Swartz, 1993). Student achievement has also been shown to be significantly related to their efficacy beliefs (Ashton & Webb, 1984; Hackett & Betz, 1989; Schunk & Swartz, 1993).

Several types of research have been conducted to measure and describe the behavior of students with different levels of self-efficacy (Riggs & Enoch, 1990; Ashton et al. 1983; Gibson & Demba 1984; Hamilton & Swortzel, 2007). Students with a strong sense of efficacy according to Gibson & Demba (1984) are more likely to challenge themselves with difficult tasks and be intrinsically motivated. These students will put forth a high degree of effort in order to meet their commitments, and attribute failure to things which are in their control, rather than blaming external factors (Pajares, 2001; Bandura, 1997 & Pajares, 2003). Self-efficacious students also recover quickly from setbacks, and ultimately are likely to achieve their personal goals (Ashton & Webb, 1984; & Hackett & Betz, 1989). Students with low self-efficacy, on the other hand, believe they cannot be successful and thus are less likely to make a concerted, extended effort and may consider challenging tasks as threats that are to be avoided (Pintrich & Schunk, 1996). Thus, according to Pajares (2001) students with poor self-efficacy have low aspirations which may result in disappointing academic performances.

Students holding low self-efficacy for accomplishing a task may avoid it while those who believe they are capable are likely to participate (Pajares, 2001).

2.8.3.3 Self-determination as a motivational component

Another important motivational component focused in this study is self-determination. Self-determination refers to a characteristic of a person that leads them to make choices and decisions based on their own preferences and interests, to monitor and regulate their own actions and to be goal-oriented and self-directing. Reeve et al. (2003) described self-determination as the ability of students to choose and control over what and how they want to learn. This motivational component has been widely propagated by Deci and Ryan (2000) in their self-determination theory (SDT). The theory posits that one's level of self-determination is determined by the satisfaction of three innate psychological needs: autonomy, competence, and relatedness. Autonomy refers to being the source of one's own behavior and achieving congruence between the activity and one's integrated sense of self. Competence refers to the need to have an effect on the environment and to achieve desired outcomes, and relatedness is the desire to feel connected to valued others (Ryan & Deci, 2002).

Self-determination theory has been widely discussed in the field of motivation in school learning, and many studies have shown that self-determination level can affect students' learning and performance and, conversely, that learning can affect the level of one´s self-determination (Eisenman, 2007; Deci & Ryan 2000; Field & Hoffman, 1994). The more these needs are satisfied, the greater the level of one's self-determination. Individuals are also more likely to pursue an activity for self-determined reasons if they feel competent because they can identify a link between their behavior and desired outcomes (Ryan & Deci, 2000). Third, one cannot function in a fully self-determined manner without a sense of volition and a feeling that the activity is concordant with one's integrated sense of self (Ryan & Deci, 2000). An advantage of this approach is that when students are given the freedom to determine their academic tasks, they are more likely to benefit from them (Glynn & Koballa, 2006).

In a study conducted by Black and Deci (2000), results obtained supported the idea that self-determination leads to improvements in student learning. They found that students with a high desire to enroll in the course were significantly correlated with perceived competence, interest/enjoyment of the course, low anxiety, and were more focused on learning whilst those who enrolled due to course requirements were significantly correlated with dropping out of the course. Field & Hoffman (2002) posited that teachers who support self-determination in students' result in a positive impact on students learning toward science and pursuing a career in science. Reeve et al. (2003) also concluded that when students believe that they have some degree of control over their learning, such as selecting some of their lab topics, overall motivation is increased.

2.8.3.4 Assessment anxiety and test anxiety

Assessment anxiety and test anxiety are two common and interchangeably terms in educational studies and they have the same meaning. The terms describe a psychological condition in which people experience extreme distress and anxiety in testing situations. Olatoye & Afuwape (2003) defined test anxiety as a psychological state of mind where a student expresses levels of worry, fear, uncertainty, concern, and helplessness before, during, or after a test. Liebert & Morris (1967) originally attributed test anxiety to two main psychological components; worry and emotionality. Worry according to Liebert and Morris (1967) refers to cognitive factors, such as negative expectations or feelings of inadequacy, and emotionality refers to the physical symptoms, such as increased heart rate, muscle tension, or butterflies. Both are aversive elements that can create anxiety, but it is the cognitive factors that have the strongest connection to performance (Liebert & Morris, 1967).

Apart from worry and tension, student's metacognitive beliefs also play an important role in the maintenance of negative self-beliefs. As a result, individuals who are test-anxious become more obsessed with the implications and consequences of failure to meet situational challenges rather than rationally focusing on completing the task in an orderly manner (Karteroliotis & Gill,1987). In addition, according to Zeidner & Schleyer (1999) test-anxious students are attacked with the feelings of inadequacy, helplessness, and anticipations of punishment or loss of status and esteem manifest anxiety responses. Similarly, it has been reported in another study that the thoughts of failure disappointing the person who motivates them may also increase test anxiety (Olatoye and Afuwape, 2003). In the same vein Putwain (2007) found that a low academic self-concept was associated with higher worry and tension about their abilities to do well on a test. Students with high expectations and thoughts of perfection face anxiety as well. They see the first position as so significant that coming in second place is considered as a failure (Morris et al. 1981).

Many motivational studies have found that test anxiety involves many negative effects including poor performance, low motivation, negative self-evaluation beliefs, and low concentration, as well as an increase in school dropout rates and general anxiety (Ben-Zur & Zeidner,1989; Morris et al. 1981). The effect of test anxiety on motivation can also influence the success expectancy. For example Ben-Zur & Zeidner (1989) and Ma (1999) found that students with high anxiety often show low confidence on their ability to cope with academic situations because they do not have the skills to cope, thus, they do not have control or are losing control of what they are doing. Consequences of failing test, unable to finish test or being embarrassed due to low grades are some similar thoughts that run through highly test anxious students' minds (Karteroliotis & Gill, 1987). Anxiety reactions can be generalized from previous experiences to testing situations. Sarason et al. (1990) also reported in their study that the level of perceived preparedness, self-efficacy, previous exposure to course materials and test anxiety significantly predicted students' achievement in science.

2.8.3.5 Content relevance as a component of motivation

Another significant component of students' motivation according to Glynn and Koballa (2006) is their perception of relevance. Frymier and Shulman (1995) define content relevance as a student perception that course content satisfies his/her interests, personal and/or career goals. As a motivational construct, content relevance can be traced to Keller's (1983) theory of motivation called Attention, Relevance, Confidence, and Satisfaction (ARCS). From this perspective, students will be internally motivated to learn if their attention is captured in class, they view the class content as personally relevant, they are confident in their ability to learn and use the course material and they feel satisfied as learners in the class. The argument towards the relationship between content relevance and student motivation put forward by Keller (1983) closely resembles with Bloom´s (1956) taxonomy of affective learning. According to Bloom (1956), students must first be willing to receive and respond to new information, experience satisfaction from engagement with this information and recognize its value before achieving the highest levels of affective response. Levitt (2001) interpreted relevance as importance, usefulness, or meaningfulness to the needs of the students. According to Osborne & Collins (2001) when students themselves decide on the topics of interest in school science, relevance takes on a personal meaning.

There a large quantity of studies describing the relationship between content relevance and students' state of motivation (Frymier, 2003; Frymier and Shulman, 1995; Webster et al. 2011). Conclusions from these studies prompted, Webster et al. (2011) to suggest that school science will only engage students in meaningful learning if the curriculum has personal value and enriches students' cultural self-identities. Both Bollinger et al. (2010) and Frymier and Shulman (1995) also found a correlation between relevance and motivation in web-based learning. According to Barmby et al. (2008) students perceive science education as relevant to them through three areas. Firstly, the usefulness of science in the society which means they are more interested to learn if the content is related to societal issues. Secondly, students' interest towards science learning which means that students are motivated to learn and do the tasks and activities in science. Lastly, the importance of science in the course they are taking which means the science content learned is meaningful and useful to them (Barmby et al. 2008)

2.8.4 Self-concept motivation and its implication to science achievement

2.8.4.1 The Concept and dimensions of self-concept

Self-concept is our perception or image of our abilities and our uniqueness. Several science and psychology researchers have studied extensively the concept of self and its connection to motivation and academic achievement (Reese et al. 2007; Bellmore, & Cillessen, 2006; Brendgen, 2002; Chapman et al. 2000; Fleming, 1984; Gans et al. 2003; Marsh et al. 2005; Marsh et al. 1988; Trautwein, et al. 2006). Marsh et al. (2005) defines self-concept as a collection of beliefs about one's own nature, unique qualities, typical behavior and self-perceptions. It is a multidimensional construct that refers to a person's perception of self in terms of both academic and nonacademic aspects (Brendgen, 2002; Chapman et al. 2000; Fleming, 1984; Gans et al. 2003; Marsh et al. 2005; Marsh et al. 1988; Trautwein et al. 2006). According to Marsh et al. (1988) the self of an individual consists of attributes and personality traits that differentiate one from other individuals. Trautwein et al. (2006) provided the meaning of self-concept as the individual's belief about himself or herself, including the person's attributes and who and what the self is. Academic self-concept, on the other hand, refers to a person's perception of self with respect to achievement in school (Chapman et al. 2000). In particular, a person's science self-concept refers to the perception or belief in his or her ability to do well in science or confidence in learning science (Marsh et al. 2005; Reese et al. 2007; Bellmore & Cillessen, 2006; Brendgen, 2002).

According to humanist psychologists, such as Carl Rogers and Abraham Maslow, humans have an inherent drive to know and express the self, resulting in the development of a self-concept (an idea of who they are) and an ideal self (an idea of who they want to be). Self- concept is the cognitive or thinking aspect of self that is self-image related (Bellmore & Cillessen, 2006). It is also the totality of a complex, organized and dynamic system of learned beliefs, attitudes, and opinions that each person holds to be true about his or her personal existence. These perceptions are influenced by a number of factors such as evaluations of significant others, reinforcements, and attributions of behavior (Marsh et al. 1988; Fleming, 1984; Marsh et al. 2005). It is the product of one’s reflectivity; it is the concept of the individual of himself as a physical, social and moral and existing being (Brendgen, 2002).

According to a theory known as social identity theory, self-concept is composed of two key parts: personal identity and social identity (Turner & Oakes, 1986). Our personal identity includes such things as personality traits and other characteristics that make each person unique. Social identity includes the groups we belong to including our community, religion, college, and other groups. A person’s self-concept is composed of evolving subjective conscious and unconscious self-assessments. Physical attributes, occupation, knowledge, and abilities of the person will change throughout the life span, contributing to changes in one’s self-concept (Turner & Oakes, 1986). In the self concept-based model of motivation, one's concept of self is composed of four interrelated self-perceptions: the perceived self, the ideal self, one's self-esteem, and a set of social identities. Each of these elements plays a crucial role in understanding how the self-concept relates to energizing, directing and sustaining organizational behavior(Turner & Oakes, 1986).

2.8.4.2 Self-concept and its connection to science academic achievement

The idea that there is a relationship between self-concept and school performance is not a new one. Self-concept and self-esteem are important factors influencing behavior and achievement in school, students tend to behave in accordance with their beliefs about themselves (Byrne, 1990). Those who consider themselves “good students” according to Marsh et al. (2005) tend to pay more attention, follow directions in class in a better way, use effective learning strategies are more likely to work independently and tirelessly to solve difficult problems, and often get enrolled in challenging courses. On the other hand, those who believe they are “poor students” misbehave in class, study rarely or not at all, abandon to turn in their homework assignments and mostly avoid taking difficult subjects (Marsh et al. 2005). Chapman et al. (2000) observed that there is a persistent and significant relationship between the self-concept and academic achievement and that change in one seems to be associated with a change in the other. Dambudzo (2005) for example conducted a study about the relationship between learner self-concept and achievement in secondary schools in Zimbabwe. The sample consisted of 1281 adolescent learners in urban and rural government and nongovernmental secondary schools. A positive and reciprocal relationship between learner self-concept and academic achievement was found.

Bellmore & Cillessen (2006) & Brendgen (2002) on their research concluded that self-concept is the basis for all motivated behavior because it gives rise to possible selves and it is possible selves that create the motivation for behavior. The academic self-concept is one aspect of self-concept because it relates to how well students do in school or how well students learn. Byrne (1990) and Reese et al. (2007) showed that academic self-concept was more effective than academic achievement in differentiating between low-track and high-track students. Self-concept encourages students or learners to develop in the study of Basic Science and this will provide necessary information for provoking an inquisitive spirit of secondary school students. Haque and Sarwat (1998) conducted a study using Academic Self-Concept Scale to investigate the age, gender and achievement effects on the academic self-concept of high school children. The results showed that there was a strong positive relationship between achievement and academic self-concept.

Hamachek (1995) found a significant correlation between self-concept and academic achievement. Ever since then, several studies have been conducted with most finding a significant correlation between academic achievement and self-concept., thereby emphasizing the importance of the self-concept for academic achievement. Hamachek (1995), following a review of self-concept literature, came to the conclusion that a relationship existed between self-concept and academic achievement and that the relationship was reciprocal, with each variable affecting the other. Consequently, learners have to do well in school in order to have a positive self-concept about their academic abilities and a positive self-concept was a necessary pre-requisite for doing well in school (Hamachek, 1995).

From these results, she also concluded that learners with a high self-concept tended to approach school related tasks with confidence and that success in those tasks reinforced their self-confidence. The opposite was also likely to be true for children with a sense of inferiority or low academic self-concept. Consequently, educators have to be sensitive to learners' self-concepts and their perceived academic achievement. Hamachek (1995) further underscores both the importance of self-concept for academic achievement and the reciprocal relationship, with the following conclusion of the review regarding self-concept research. According to Hamachek (1995) it is difficult to find ways to help students do better in school without also exploring ways to help them feel better about themselves as learners. At the same time, it is almost impossible to help students improve their self-concepts without assisting them in finding ways to improve their school performance (Hamachek, 1995).

Interestingly however, Marsh et al. (2005) argued that reciprocal effects models of longitudinal data show that academic self-concept is both a cause and an effect of achievement. In support of this view, Afuwape (2011) presented findings that suggested a negative relationship between self-concept and achievement, noting that students have the highest academic performance and achievement had some of the lowest overall beliefs in their perceived self-concept. In their longitudinal study on academic self-concept, interest, grades, and standardized test scores Marsh et al. (2005) found that despite stereotypic gender differences in means, linkages relating these constructs were invariant over gender. Their results demonstrated the positive effects of academic self-concept on a variety of academic outcomes and integrate self-concept with the developmental motivation literature. In conclusion, these respective findings imply that the relationship between self-concept and achievement differs depending on the context or, methodologically speaking, the unit of analysis.

2.9 Literature gap

Despite numerous studies in the area of science process skills and test development (Dillashaw & Okey, 1980; Padilla et al. 1983; Roth & Roychoudhury, 1993; Arena, 1996; Harlen, 2000; Dyer et al. 2004; Lambda & Fraiser, 2004; Mungandi 2005; Hamilton & Swortzel, 2007), none of them had been conducted in Tanzania or comparable countries. Studies by Mushi, 1992; Chonjo et al. 1995; Osaki, 2000; Osaki & Njabili, 2004; Osaki, 2007; Shemwelekwa, 2008; Semali & Mehta, 2012) focused much on how science was taught in Tanzania schools and the availability of science teaching and learning resources. A review of the literature failed to identify any study in Tanzania that have investigated whether or not Tanzania students are acquiring competence in science process skills as prescribed in the new curriculum. The current researcher was convinced that there is an educational gap in this area.

A review of literature further failed to identify any study that had attempted to construct and validate a scientific test for measuring science process skills in the context of Tanzania. On the other hand, most of the available developed tests for science process skills have nonsubject specific questions. They tend are made up of Biology, Chemistry and Physics questions altogether (see test of integrated process skills by Dillashaw and Okey (1980), integrated science process skills test TIPS II by Burns et al. (1985), process skills test by McKenzie and Padilla (1986), science process skills test by Onwu and Mozube (1992), integrated process skill by Mungandi, 2005 and the recent one by Shahali & Halim (2010). However, as Millar & Driver (1987) believe, scientific inquiry cannot be independent of domain knowledge. Therefore with the recent science curricular reforms in Tanzania and the absence of a valid measuring tool, it becomes vital for this study to come up with a valid tool to be used in examining students' level of science process skills. The tool that would be used to measure whether or not students have acquired prescribed science process skills even in the absence of well-equipped science laboratories and in the overcrowded classes.

Despite numerous studies on the value of inquiry teaching approach worldwide and its acknowledgment in the Tanzania syllabuses, review of literature and studies failed to identify any study that scientifically investigated the effectiveness of the approach on students' scientific process skills development, conceptual understanding of Biology contents and motivation. On the other hand, a review of learning motivation studies revealed the diversity and variety of motivation factors, such as self-perceptions of ability, intrinsic goal orientation, self-efficacy, grade motivation, self-regulated learning, task orientation and self-determination (Glynn & Koballa 2006; Eccles & Wigfield 2002; Schunk et al. 2008; Pintrich & Schunk, 1996). However, no study have been conducted in Tanzania to find out the influence of inquiry-based approach to the development of students´ intrinsic motivation, extrinsic motivation, self-efficacy, self-determination or perception of science content relevance components as a single motivational entity.

CHAPTER THREE
GENERAL METHODOLOGICAL ISSUES

3.1 Introduction

This part describes the general methodological issues that were across at different stages of this study. As it has been described in the earlier sections that this study on science process skills was conducted stage wise. The first stage was the development and validation of a test for measuring students' knowledge level of integrated science process skills specific to Biology. The second stage employed the test that has been developed and validated in the first stage in measuring and assessing the knowledge level of science process skills of advanced level Biology students in the municipality of Morogoro Tanzania. The last stage was a quasi-experimental study to investigate the effectiveness of inquiry-based teaching approach on the development of students' scientific process skills, conceptual understanding of genetics and motivation towards science process skills. This part just describes the general methodological issues that were across all these three stages of this study. Specific study methodologies for each stage have been presented with the corresponding discussion of findings in next chapters.

Hence this part spells out only the general methodological aspects such as research designs of the whole study, the area of the study, population, and some ethical issues that were considered while undertaking the study. Specific issues of like sampling procedures, sample size, data collection tools, data analysis techniques depended on the stage in question and are presented in the chapters of specific research methodology sections. Because of the nature of the study itself, quantitative techniques were the dominant research paradigm adopted.

3.2 Area of the study

This study (all stages of the research) was conducted in Morogoro municipality Tanzania. The area was conveniently selected by the researcher to represent other regions of the country. The basis for convenience selection of Morogoro municipality stems from the fact that, advanced level Tanzania Biology students throughout the country are undergoing a uniform centralized competence based curriculum. Moreover, these students have been selected to pursue a higher secondary education in Biology based on their good performance in their National Form IV examination results in the subject. The assumption here is that all Tanzania Biology students are somewhat similar in terms of their science process knowledge and skills. In the case like this, where the researcher has a wider possibility of obtaining the needed data, he/she is allowed to search a sample by his/her convenience (Borg and Gall, 1989). Morogoro municipality is a convenient area for the researcher because it is the researcher's working station making easier for him to obtain enough needed data.

3.3 Description of Morogoro Municipality

Morogoro municipality is the regional headquarters of Morogoro. It covers a total area of 531square kilometers with 29 administrative wards. It is located about 195 kilometers to the west of Dar es Salaam (see fig 3.1 below) and is situated on the lower slopes of Uluguru mountains whose peak is about 1,600 feet above sea level. It lies at the crossings of longitudes 37.0 east of the greenwich meridian and latitude 4.49 south of equator. According to the 2012 census the current population of the municipality was 315,866 of which 164,166 (52.15%) are women and 151,170 or 47.85% are men. The major physical features include the famous Uluguru mountains, which lie in the southeastern part, and Mindu mountains, which lie in the western part. There are three main rivers with several tributaries, which form a number of alluvial flood plains. These rivers are the Morogoro, Kilakala, and Bigwa. Other sources of water are the Mindu Dam' which serve for the industrial activities as well as domestic purposes.

Despite the variation of climatic conditions throughout the year, Morogoro municipality experiences an average temperature of 30oC degrees centigrade with a daily range of about 5oC degrees centigrade below or above 30 oC. The highest temperature occurs in November and December (mean of 33 oC), and the minimum temperature is in June and August when the temperatures go down to about 16oC degrees centigrade. The mean relative humidity is about 66% and drops down to as far as 37%. The total average annual rainfall ranges between 821mm to 1,505mm. Long rains occur between March and May and short rains occur between October and December each year and the average monthly amount of precipitation is at around 12 mm.

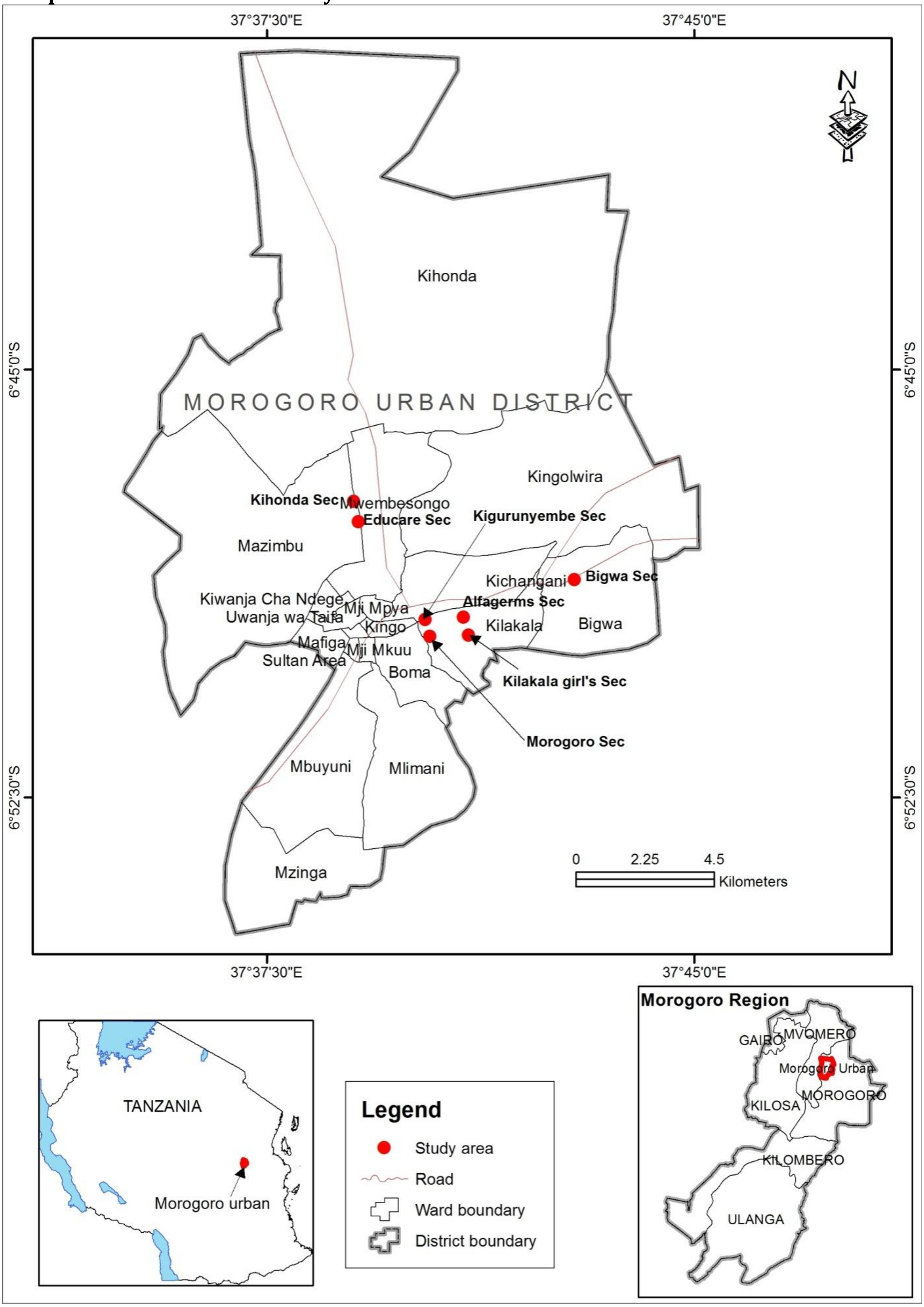
The key education sector institutions are pre-primary classes, primary schools, secondary schools, vocational centers, specialized training centers, collages and a university as shown in Table 3.1 below. The special education offer education to the disabled children particularly the blind, the dumb the deaf and autism. The municipality of Morogoro in particular has a total of 46 secondary schools, of which 23 are government schools and 23 are privately owned (Municipal Council Data, 2014).

Table 3.1 Education Institutions found in Morogoro municipality

EDUCATION FACILITY		OWNER	
CATEGORY	**No. OF FACILITIES**	**GOVERNMENT**	**NON - GOV**
Pre-Primary schools	61	38	23
Primary schools	85	62	23
Secondary schools	46	23	23
Vocational Centers	03	02	01
High Institution (Universities)	03	01	02
Special education for disabled	11	10	01
Teacher Trainin-g Collage	1	1	0

Source: Municipal Education Office (2014)

Figure 3.1 The map of Morogoro municipality showing the location of the sampled schools for this study



3.3 Study population

The population for this study was the advanced level Biology students (Form V and VI) who have Biology as one of their major subjects present in the municipality of Morogoro. However, in the first stage which intended to develop and validate the process skills test, the researcher also used the ordinary level students (Form IV) in order to obtain data for calculating construct validity of the constructed test. The researcher decided to involve advanced level students because they are higher level learners where integrated science process skills such as identifying experimental questions, identifying variables, formulating hypotheses, designing investigations, graphing and interpreting data are a vital aspect of their meaningful science activity (Mattheis & Nakayama, 1988). Advanced level students are expected to have acquired competence in the integrated process skills as planned in the curriculum. These process skills represent the rational and logical thinking skills required in the process of problem-solving in science. According to the basic educational statistics of Tanzania (2014), Morogoro municipality has a total 1880 advanced level students of which 784 are female students and 1096 are male. Three hundred and fifty-three (353) potential participants (Form V and Form VI students) have been identified from the current list of Physics, Chemistry, and Biology (PCB) students and Chemistry, Biology and Geography (CBG) students from different schools of present in the municipality of Morogoro as provided by the office of the district education officer, division of secondary education.

3.4 General research design

The study was conducted in three stages. The first stage developed and validated the test of integrated science process skills specific to Biology (BPST). This paper-and-pencil objective test was developed specific to the Biology contents as defined in the Tanzania Biology syllabus for advanced level students. The development of BPST was based on the school of thought that scientific inquiry cannot be completely independent of knowledge. The test specifically measured students' performance on specific integrated science process skills which includes skills in i. formulating hypotheses, ii. defining variable operationally, iii. identifying and controlling variables, iv. planning investigations as well as v.

analyzing and interpreting data. Evidence of content validity, construct validity, internal consistency reliability, difficult index, and discrimination index was investigated to prove its psychometric properties.

The second stage of the study employed the test that has been validated test in the first stage (BPST) to examine the knowledge level of science process skills of advanced level secondary school Biology students in Morogoro municipality. This included examining the performance of students in individual integrated science process skills (formulating hypotheses, defining variable operationally, identifying and controlling variables, planning investigations and analyzing and interpreting data) and determining whether there was statistical difference in performance of students based on individual science process skill. This study at this stage was conducted in order to establish a base level of information on whether or not Tanzania students are acquiring the prescribed competences present in the competence based curriculum of 2005. It has to be noted that in 2005 Tanzania came up with the so called 'Competence Based Curriculum' which emphasized among other things, student's competence in science process skills. The curriculum emphasized the need of Tanzania science Students to learn scientific subjects such as Biology, Physics and Chemistry in the same way science is done scientists. Despite such a dramatic shift in curriculum policy, little is known about whether or not the reform efforts are truly transforming the educational experiences of students.

The third stage investigated the effectiveness of inquiry-based teaching approach on students' scientific process skills development, conceptual understanding of Biology contents (case study being genetics) and motivation towards science process skills. Eight (08) weeks genetics lessons were designed for intervention from Tanzania Biology syllabus on the basis of both inquiry based learning principles and conventional lecture style. The intervention was implemented into Form V and VI Biology students from 03 selected schools in the Municipality of Morogoro. Throughout the teaching of genetics in the experimental groups, Biology students worked in small groups where they were encouraged to explore problems, discuss, formulate hypotheses, share their ideas with their classmates, discuss their observations and interpret findings. In the control group, students learned genetics conventionally through prescribed books and lecture notes. A quasi-experimental research design with a pre-test and post-test was used in this stage. Students completed the same data collection instruments before and after instruction/intervention so that changes in their conceptual understanding, scientific process skills and motivational constructs (intrinsic motivation, grade motivation, career motivation, self-efficacy, self-determination and self-concept) towards science process skills can be spotted. The validated test of science process skills (BPST) was used to measure students' performance in science process skills, Genetics test on the other hand was used in measuring students conceptual understanding (Genetics being the case study), the Science Motivation Questionnaire II (SMQ-II) developed by Glynn et al. (2011) and FSWEx questionnaire developed by Damerau (2012) were employed for measuring students' level of motivation (intrinsic motivation, grade motivation, career motivation, self-efficacy, and self-determination) and self-concept towards science process skills respectively before and after intervention.

Figure 3.2 Diagramatic representation of the adopted research design for the study presented in the following chapters

Stage one: Development and validation of Biology process skills test (BPST)

- Item collection modifications and moderations
- Conducting pilot study in order to generate data to be used in determining validity, reliability and item effectiveness
- Determining the content validity, construct validity, internal consistency reliability, difficult indices, and discrimination indices of items in the test (BPST)

⇓

Stage two: Examining students knowledge level of science process skills

- Determining general performance in the Biology Process skills test
- Determining the knowledge level of science process skills based on sex and grade level of students
- Assessing the performance in each process skill understudy i. formulating hypotheses, ii.defining variable operationally, iii.identifying and controlling variables, iv. experiment design and v. interpreting data

⇓

Stage three: Quasi-experimental study to determine the effectiveness of inquiry-based teaching approach in

- Students' scientific process skills development,
- Students' conceptual understanding of Biology contents , genetics being the case study
- Students' motivation towards science process skills
- Determining the correlation between the knowledge of science process skills with conceptual understanding, and motivation constructs

CHAPTER FOUR

STAGE ONE: METHODOLOGY, RESULTS, AND DISCUSSION OF THE STUDY TO DEVELOP AND VALIDATE THE SCIENCE PROCESS SKILLS TEST (BPST)

4.1 Introduction and rationale of developing an integrated science process skills test

As described in the earlier sections, the first stage of this study aimed at developing and validating science process skills test specific to Biology (BPST). This paper-and-pencil objective test was developed specific to the Biology contents as defined in the Tanzania Biology syllabus for advanced level students. This study was based on the school of thought that scientific inquiry cannot be completely independent of knowledge. The test specifically measured students' performance on specific integrated science process skills which includes skills in i. formulating hypotheses, ii. defining variable operationally, iii. identifying and controlling variables, iv. planning investigations as well as v. analyzing and interpreting data. The test consists of 35 multiple-choice items and took into account realities of the Tanzania education system. Evidence of content validity, construct validity, internal consistency reliability, difficult index, and discrimination index was investigated to prove its psychometric properties.

The science educational reforms of the 1960’s and 1970’s prompted the need to develop various test instruments for assessing both teachers and students’ level of science process skills (Dillashaw & Okey, 1980 & Mungandi, 2005). Since then, scientific process skills evaluations performed through paper and pencil multiple-choice test have been very famous (Burns et al. 1985; Dillashaw & Okey, 1980). However with the recent science education reforms there has been a considerable reaction to and criticism towards traditional paper and pencil tests type of assessment in assessing the performance of tasks. These critics proposed that students can authentically be assessed in their ability to perform a task, like the science process skills when they perform the real task. According to Wiggins (1989), authentic assessment means that tests should involve real-life tasks, performances, or challenges that replicate the problems faced by a scientist, historian, or expert in a particular field. Authentic tests are complex tasks rather than drills, worksheets, or isolated questions. Harlen (1999) for example asserted that science process skills were inseparable in practice from the conceptual understanding involved in learning and applying science and played a central role in learning with understanding. Specifically in the context of student performance in school science laboratory, Gronlund (1998) wrote that "if you want to determine if students can conduct an experiment, have them conduct an experiment".

However, authentic methods for assessing science process skills competence such as through laboratory practical work have a number of constraints particularly in the context of teaching in large under-resourced science classes (Onwu & Stoffels, 2005 & Mungandi, 2005). This is particularly true also regarding Tanzania secondary schools. A survey conducted by Osaki (2007) Osaki & Njabili (2004), Semali & Mehta (2012) & Athuman (2010) on the teaching of science in selected schools in Tanzania reveals science classes being characterized by a large number of students. This makes impossible for effective guiding of science practicals and for authentical assessment of students. Moreover, most of these schools, especially community owned secondary schools either do not have science laboratories or they are poorly equipped with of reagents, apparatus, and samples (Athuman, 2010 & Osaki, and Njabili, 2004). Objective tests in the multiple choices format is an alternative for measuring students´knowledge of science process skills where authentic assessment of these skills is impossible. This is the common way that has been used worldwide to assess science process skills, especially in large under-resourced science classes. Multiple choice tests through the use of paper and pencil test do not require laboratories and expensive resources. Hofstein & Lunetta (2004) for example suggest the use of both authentic and alternative assessment methods in measuring outcomes of school science lab programs. The authors asserted that even today in an era of highly emphasised standards approach to science, assessment of students' performance in the science labs should be confined to both conventional using objective, pencil and paper and authentically.

4.2 Methodology in the development of a science process skills test (BPST)

This part of methodology presents procedures taken in constructing and then validating an integrated science process skills test specific for Biology (BPST). Specifically, the section spells out i. research design adopted ii. data collection instruments used iii. the procedure adopted in test construction and validation, iii. piloting of the test, iv. ethical issues, and vi. analysis of the pilot study data.

4.2.1 Research design

This stage of the study was basically instrumentation. It was aimed at construction, validation and production of valid and reliable test for assessing Biology students' knowledge of science process skills. Quantitative survey type research design was adopted. Selected schools in Morogoro were surveyed and the developed test administered. Quantitative data from the pilot study was used in the determination of its psychometric properties (reliability, validity, and items effectiveness). All the items were biased to Biology discipline and guided by the Tanzania advanced level Biology syllabus which is a product of competence based curriculum of 2005. The test constructed is now called Biology process skills test (BPST).

4.2.2 Data collection instruments

Three different kinds of instruments were used in this stage. They include,

i. The constructed Biology process skills test (BPST). This instrument was developed by the researcher for the purpose of validating it. The instrument was used for collecting data which was then used in determining its quality in terms of validity and reliability and item effectiveness. Students' score from this test in during pilot study were used. The instrument also provided data for comparison of students' performance in different process skills under question (see appendix I).

ii. The test of integrated science process skills (TIPSII). This instrument (TIPSII) was developed by Burns et al. (1985). In this study, this test was used in the determination of the concurrent validity of the developed Biology process skills test (BPST) through comparison of student scores (see appendix II).

iii. Content validation form. This tool was prepared to guide the test experts in determining the content and face validity of the developed BPST. This tool is a 3 point scale which asked raters opinions on whether or not the item had met or not met a certain criterion quality. The degree of congruence between rates was taken as an evidence of face and content validity of the developed BPST (see appendix III).

4.2.3 Procedures in development and validation of a test (BPST)

According to Burton & Mazerolle (2011), instrument development includes four steps. Step one consists of defining constructs and determining domain content. Step two involves generating items for the survey and judging the appropriateness of the items. Step three is to design and conduct studies to test the scale. Lastly, step four involves finalizing the scale based on data collected in the third step. To pilot test an instrument, researchers must consider sample size, sample composition, initial item reliability estimates, and the type of validity-related surveys to include in the study's design (Burton & Mazerolle, 2011). However, in this study, the following steps as suggested Spector (1992) were followed throughout the process of developing and validating the Biology process skills test (BPST).

i. deciding on the principles upon which the constructed test will be based
ii. defining constructs and objectives to be measured
iii. collection and preparation of test items
iv. conducting initial validation
v. carrying out pilot testing
vi. performing item analysis,
vii. establishing reliability evidence
viii. calculating the validity of test scores
ix. estimating readability index of the test, and finally
x. preparing the final valid test.

4.2.3.1 Principles that guided BPST construction

Several principles guided the researcher in constructing and validating this science process skills test. The intention was to produce a test having psychometric characteristics within the acceptable statistical ranges. The following are the characteristics of the test that the study wanted to produce.

i. The test that will be made up of items referenced to specific integrated science process related to 1) identifying and controlling variables, 2) defining operationally, 3) formulating hypotheses, 4) experimenting design and 5) analysis and interpretation of data
ii. The test that had to be subjected to rigorous procedures of determining item analysis, validity, readability and reliability measures so that it is good enough for measuring identified skills.
iii. The test that will have items biased to Biology discipline and guided by the Tanzania advanced level Biology syllabus (a product of competence based curriculum of 2005).
iv. The test which will be a multiple choice type, in which each process skill had to be measured by seven (07) items making a total of 35 questions so that a student demonstrate competency for each individual process skill
v. The test that will have a moderate length such that a majority of student complete within one hour (60 minutes) of administration
vi. The test that will be called Biology process skills test (BPST)

4.2.3.2 Defining constructs and objectives to be measured by BPST

After defining the principles that governed the test to be constructed, the second stage in the test development was to define constructs and objectives to be measured. However, the essential aspects of science process skills as a construct had already been determined by the researcher through reviewing literature and theories (see section 1.9). Identifying the science process skills constructs was done in order to guide the researcher define them clearly and precisely. Integrated science process skills identified included 1) identifying and controlling variables, 2) defining operationally, 3) formulating hypotheses, 4) experimenting skills, and 5) interpreting data and drawing conclusions (Chiappetta & Koballa, 2002). On the other hand, the objectives to which test items were referenced was based on Dilashaw and Okey's (1980) and Mungandi (2005) tests as shown in table 4.1 below.

Table 4.1: Objectives upon which BPST test items was based

Process skill measured	Objectives
Identifying and controlling variables	Given a description of an investigation, identify the dependent and independent variables
Operational definitions	Given a description of an investigation, identify how variables are operationally defined
Identifying and controlling variables	Given a problem with a dependent and independent variable specified, identify the variables which may affect it
Stating hypothesis	Given a problem with dependent variables and a list of possible independent variables, identify testable hypothesis
Operational definitions	Given a verbally described variable select a suitable operational definition for it
Stating hypothesis	Given a problem with a dependent variable specified. Identify a testable hypothesis
Designing investigations	Given a hypothesis, select a suitable design for an investigation to test it
Graphing and interpreting data	Given a description of an investigation and obtained result or data, identify a graph that represent the data
Graphing and interpreting data	Given a graph or table of data from an investigation, identify the relationship between variables

Adopted from test of integrated science process skills for secondary school developed by Dillashow, F.G and Okey, J.R (1980)

4.2.3.3 Item collection and preparation

Items for BPST instrument were collected based on the constructs and objectives stated above. The major efforts in item preparation were focused on researchers own experience and refinements or modification of items from other existing integrated science process skills tests. That means questions were initially collected from various sources such as local science past papers, science achievement tests, textbooks and related science process skills test. Potential scientific process skills test such as the test of integrated process skills by Dillashaw and Okey (1980), the group test of integrated process skills by Tobin and Capie (1982), integrated process skills test TIPS II by Burns et al. (1985), a test of integrated science process skills by Mungandi (2005), the performance process skills test for middle grades (POPS) by Mattheis & Nakayama (1988), and the science process skills test by Onwu & Mozube (1992) were targeted. Table of specifications which is a blue print in test construction were used to ensure that each selected process skill has a place in the test. Although the aim was to develop a 35 items test, a total of 43 items from various sources were collected, constructed and moderated ready for initial validation.

4.2.3.4 Initial BPST validation (ensuring face and content validity)

After collection of potential items and conducting test moderation, the test was subjected to analysis of face and content validities. This was an initial validation activity which was carried out by experts before piloting the tool to a smaller sample in Tanzania. According to Anastasi & Urbina (1997), content validity is a non-statistical type of validity that involves the systematic examination of the test content to determine whether it covers a representative sample of the behavior domain to be measured. Face validity, on the other hand, is the extent to which a test is subjectively viewed as covering the concept it purports to measure. In other words, a test can be said to have face validity if it "looks like" it is going to measure what it is supposed to measure by observant.

In this study, content validity, as well as face validity of the constructed test, was assured by using a panel of experts who reviewed the test specification and the items selected. The experts reviewed items and commented on whether the items cover a representative sample of the behavior domain. Content validation form was prepared as a guide for the experts (see appendix III). The experts also provided answers to the test items so as to verify the accuracy and objectivity of the scoring key. Questions that raters were not satisfied or found to have serious flaws were either modified or discarded. Six science educators (two from the University of Wuppertal and 04 from higher secondary schools in Morogoro) with experience in both test construction and the science process skills reviewed the draft. The raters reviewed content, quality in relation to the specific content domain. They also corrected all grammatical problems and flagged problematic items. The concurrence of raters in the form was taken as evidence of content validity and objectivity of scoring. All the comments from experts were used to revise and improve the test questions. This is was the first validation and it resulted into discarding some items from 43 to the required 35.

4.2.3.5 Pretesting the constructed process skills test (Pilot study)

After non-statistical validation of the test by the experts, the test was then subjected to pilot study in the selected schools in Morogoro. Pre-testing is the administration of the data collection instrument with a small set of respondents from the population before full-scale application. The pilot study generated data which was then used in determining validity and reliability of the test as well as assessing the effectiveness of items in the test. The subjects involved in the pilot study with the developed test were the advanced level Biology students (Form, V, and VI) who have Biology as one of their major subjects. Form IV students from convenience schools were also used for the purpose of obtaining data needed in calculating construct validity of BPST.

Schools in which Form IV students participated in the pilot study were selected by convenience and their willingness to participate. Pilot testing took one month and was strictly voluntary. The identity of all students participating remained confidential. The only demographics collected were gender and grade level. The scores were not reported back to teachers or students and will only be used for the purpose of analysis. The test papers were collected at the end of each pilot and the teachers were not allowed to keep or make copies.

4.2.3.6 Performing item analysis of the BPST test

After the pilot study, the next stage in test development was to perform item analysis. The data generated during the pilot study were used in performing item analysis. Item analysis is a process which examines student responses to individual test items in order to assess the quality each item and of the test as a whole. Item analysis is especially valuable in improving items which will be used again in later tests, but it can also be used to eliminate ambiguous or misleading items in a test. In this study, item analysis included three statistics i. difficulty index, ii. discrimination index and distracters analysis. Difficulty indices which range from 0.0 to 1.0 were obtained by calculating the percentage of students who selected the correct response. An item having difficulty indices between 0.2 - 0.8 were retained and all other items out of this range were either discarded or modified (see Nitko, 1996).

Discrimination indices of items were calculated by computing the difference of difficulty indices of higher (27%) achievers and lower (27%) achievers in each item. Discrimination index indicates how well the question separates students who know the material well from those who don't. In this study, items having the discrimination index of 0.3 and above were described as good items and hence were retained for the final test (Adkins, 1974; Hinkle et al. 2003).

In addition to examining the performance of test items as a whole, the study also examined the performance of individual distracters, (answer options or alternatives). This was done by calculating the proportion of students who choose each answer option hence identifying which distracters are "working" and appear attractive to students and which ones are not. Distracters that had poor or negative discrimination index were either improved or replaced. All distracters having indices between 5% to 30% were qualified and retained in the final draft of the test (see Brown, 2000).

4.2.3.7 Estimating reliability and standard error of measurement of BPST

Item analysis was followed by estimation of its reliability and standard error of measurement (SEM) of BPST instrument. Pilot study data were also used for this purpose. Reliability is a measure of consistency with which the instrument (questionnaire, test or examination) produces the same results under different but comparable conditions. Reliability is one of the very important quality aspects of a measuring instrument. According to Braun (1988), a test should be sufficiently reliable to permit stable estimates of the ability levels of individuals in the target group. Although several reliability measures exist such as test-retest reliability, inter- rater reliability, and alternative form reliability, in this study the evidence of the reliability of BPST was obtained by calculating its internal consistency reliability. Internal consistency reliability is an easy way of obtaining reliability evidence as it involves only one test administration. The internal consistency coefficient indicates the extent to which all the items are measuring the same ability or trait.

Cronbach's alpha statistical method was used for estimating the internal consistency of a constructed test in this study. As recommended by several statisticians (Hinkle, at al. 2003; Adkins, 1974), the researcher in the current study was satisfied with reliability coefficients from 0.7 and above to have the reliability evidence of the test.

On the other hand, the pilot data was also used to determine standard error of measurement of the test scores. Since all measurement contains some error, it is highly unlikely that any test will yield the same scores for a given person each time they are retested. The standard error of measurement (SEM) estimates how repeated measures of a person on the same instrument tend to be distributed around his or her "true" score. Standard error of measurement (SEM) was calculated using the formula given below

$SEM = SD\sqrt{1-r}$

Where SEM= Standard error of measurement

SD = Standard deviation

r= reliability coefficient

4.2.3.8 Establishing validity evidence of BPST instrument

After establishing reliability evidence of the tool, the next step was to assess its validity. Validity is the extent to which a measurement procedure is capable of measuring what it is supposed to measure. Traditionally, validity is subdivided into content validity, face validity, construct validity and criterion-related (concurrent & predictive) validity (Schmidt & Hunter, 1998). As it has already been described in the earlier sections, content validity and face validity of the constructed test were already determined by a group of experts before piloting it. Apart from the analysis of content validity and face validity of the constructed test, it was necessary also to determine its concurrent validity and construct validity. Concurrent validity is the degree to which the scores on a test are related to the scores on another already established or to some other valid criterion available at the same time. In this study, the integrated process skill test (TIPS II) developed by Burns et al. (1985) was used to establish concurrent validity of BPST instrument. Advanced level Biology students in Morogoro concurrently completed Biology process skills test (BPST) and TIPS II. Concurrent validity of BPST was calculated by finding the coefficient of correlation of students score in the BPST and their scores in integrated process skill test (TIPS II).

Construct validity, on the other hand, is the degree to which a test measures an intended hypothetical construct. It involves the experimental demonstration that a test is measuring the construct it claims to be measuring (Mungandi 2005). It is determined through comparison of the performance of two groups on the test where one group is known to have the construct under the question (Mungandi, 2005). In this study, construct validity of BPST was calculated by comparing the performance of Form IV to that of advanced level students (Form V and Form VI) assuming by that the advanced level students are more knowledgeable in science process skills than the Form IV. In both cases (concurrent and construct), Pearson product moment (r) was used to find out whether there is a correlation between scores using SPSS version 21. The researcher was satisfied with validity coefficients from 0.7 and above as recommended by several statisticians (Hinkle et al. 2003; Adkins, 1974) to establish concurrent validity evidence of the test.
On the otherhand, a coefficient of correlation of less than 0.85, as proposed by Voorhees et al. (2015) was taken to conclude the existence of discriminant validity (construct validity) between the oldinary and advanced level students.

4.2.3.9 Estimating readability index of BPST instrument

Readability statistics are good predictors of the level of difficulty of documents or text, particularly technical ones. Therefore it was necessary for the study to establish readability index of BPST as one of test quality criteria. This is because the test focuses Tanzania students in which English is their second or third language. Readability is the level of ease or difficulty with which text material can be understood by a particular reader who is reading that text for a specific purpose(Kincaid et al. 1981). Both the grade level readability score and reading age readability of BPST instrument were calculated. In both cases the Flesch-Kincaid readability formulas were employed. The Flesch Reading Ease score is part of the best-known readability scores. The following Flesch steps were followed when calculating the grade level (GL) readability score

Step 1: Calculating the average number of words used per sentence.

Step 2: Calculating the average number of syllables per word.

Step 3: Multiplying the average number of words by 0.39 and add it to the average number of syllables per word multiplied by 11.8.

Step 4: Subtracting 15.59 from the result.

The ***Flesch-Kincaid Grade Level*** *readability age score* formula is:

FKRA = (0.39 x ASL) + (11.8 x ASW) - 15.59

where:

- **ASL;** average sentence length in words or average number of words in sentence (number of words divided by the number of sentences)
- **ASW;** average number of syllables per word (the number of syllables divided by the number of words)

The Flesch-Kincaid Ease Readability (RE) formula on the other hand was used to determine ease readability of the developed instrument. This implies the readability of the final test instrument. The specific mathematical formula is:

RE = 206.835 – (1.015 x ASL) – (84.6 x ASW) ,

where

RE= Readability Ease

ASL = Average Sentence Length (i.e., the number of words divided by the number of sentences)

ASW = Average number of syllables per word (i.e., the number of syllables divided by the number of words)

The Flesch Reading Ease Formula is a simple approach to assess the grade-level of the reader. It's also one of the few accurate measures around that one can rely on without too much scrutiny (Mungandi, 2005). This formula is best used on school text. It has since become a standard readability formula used worldwide. However, primarily the formula is used to assess the difficulty of a reading passage written in English (Mungandi, 2005). The following score mapping table as suggested by Mungandi (2005) was used to interpret the level of reading difficult of BPST instrument using the score obtained from the Flesch reading ease formula.

Table 4.2 Readability score mapping table by Mungandi (2005)

Score	School Level	Notes
100.00-90.00	5th grade	Very easy to read. Easily understood by an average 11-year-old student.
90.0–80.0	6th grade	Easy to read. Conversational English for consumers.
80.0–70.0	7th grade	Fairly easy to read.
70.0–60.0	8th & 9th grade	Plain English. Easily understood by 13- to 15-year-old students.
60.0–50.0	10th to 12th grade	Fairly difficult to read.
50.0–30.0	College	Difficult to read.
30.0–0.0	College Graduate	Very difficult to read. Best understood by university graduates.

According to Mungandi (2005), the score of 90.0–100.0 implies easily understood by an average 11-year-old student, 60.0–70.0 easily understood by 13- to 15-year-old students while 0.0–30.0 best understood by university graduates. The Flesch reading ease scale is rated from 0 to 100. A high readability value implies an easy to read the text. Though simple it might seem, according toKincaid et al. (1988) the Flesch Reading Ease Formula has certain ambiguities. For instance, periods, explanation points, colons and semicolons serve as sentence delimiters; each group of continuous non-blank characters with beginning and ending punctuation removed counts as a word; each vowel in a word is considered one syllable subject to: (a) -es, -ed and -e (except -le) endings are ignored; (b) words of three letters or shorter count as single syllables; and (c) consecutive vowels count as one syllable(Kincaid et al. 1988).

4.2.3.10 Preparation of the final test draft

After undertaking the above item analysis procedures and determination of test validity and reliability qualities, the final draft of the test was prepared. Poorly constructed items and destructors were either removed or modified to function in the way they were supposed. Consideration of test duration, as well as instructions for students, was taken into account once again. The final draft of now a valid and reliable Biology process skills test (BPST) was printed and photocopied confidentially ready for large scope assessment in the next stage.

4.3 PILOT STUDY WITH THE CONSTRUCTED BSPT

As is has been described in earlier sections of methodology, the developed tool was tried out to a sample of students in order to generate data which was used in determining its validity, reliability and in the assessment of the effectiveness of each individual items.

4.3.1 Sample size and schools involved in the pilot study

At least to researcher´s knowledge, there is no method available in the calculations of sample size needed to assess content validity, language validity, construct validity and internal consistency reliability in a test or questionnaire validation studies. This has also been the observation by Anthoine et al. (2014), who wrote that "sample size determination for psychometric validation studies is rarely ever justified a priori, this emphasizes the lack of clear scientifically sound recommendations on this topic". However, rule of thumb do exist which guide researchers in sample size determination. One common rule of thumb is to ensure a person-to-item ratio of 10:1. Another rule is to ensure that N = 300 (Worthington & Whittaker, 2006).

In this study, a person-to item ratio of 10:1 rule of thumb was adopted. The constructed test (BPST) is a 35 multiple choice items test. According to this rule of thumb (10:1), a sample of 350 advanced level students are sufficient. However, some researchers have criticized these rules, noting the appropriate sample size is dependent on the features of the gathered data. They recommend obtaining the largest possible sample because the adequacy of the sample size cannot be determined until after the data have been analyzed (Henson & Roberts, 2006). Hence in this study, a total of 610 students were involved instead of 350, in which 339 were Form IV and 271 were advanced level (Form V and VI). A total of seven schools (Kilakala, Afa Germs, Bigwa Sisters, Educare, Kihonda Mororgoro and Kigurunyembe) were involved. Sampling of these schools was purposive for the advanced level schools and by convenience for the ordinary or lower level secondary schools. The number of subjects in each school involved, their grade levels and sex has been shown in the table 4.3.

Table 4.3 Subjects and schools involved in the pilot study

School Name			Type of Student			Total
			Form six	Form five	Form four	
Kilakala	Sex	Female	48	91	-	139
	Total		48	91	-	139
Alfagerms	Sex	Female	22	25	-	47
		Male	25	37	-	62
	Total		47	62	-	109
Bigwa sisters	Sex	Female	9	14	-	23
	Total		9	14	-	23
Morogoro sec	Sex	Female	-	-	52	52
		Male	-	-	66	66
	Total	-		-	118	
Kihonda	Sex	Female	-	-	45	45
		Male	-	-	47	47
	Total	-		-	92	
Kigurunyembe	Sex	Female	-	-	19	19
		Male	-	-	35	35
	Total	-		-	54	
Educare	Sex	Female	-	-	39	39
		Male	-	-	36	36
	Total	-		-	75	
Total	**Sex**	**Female**	**79**	**130**	**155**	**364**
		Male	**25**	**37**	**184**	**246**
		Total	**104**	**167**	**339**	**610**

Research survey, (2014)

Pilot study also intended to generate data to be used in the determination of criterion related, the concurrent validity of the new instrument (BPST). Concurrent validity is a measure of how well a particular test correlates with a previously validated measure. In this study, a valid and reliable test of integrated science process skill (TIPS II) by Burns et al. (1985) was used determine the concurrent validity of BPST. The BPST and TIPS II were concurrently administered to among 248 advanced level students in Kilakala and Alfa Gems secondary schools. The schools were also purposefully involved. Table 4.4 summarizes the number of subjects involved in concurrent validation based on their their grade levels and sex.

Table 4.4 Subjects and schools involved in the concurrent validation of BPST during pilot study

sn	School name	Sex		Grade levels			Total
		Male	Female	Form IV	Advanced level		
					V	VI	
1	Kilakala sec school	-	139	-	91	48	**139**
2	Alfa Germs sec school	62	47	-	64	45	**109**
	Total	62	186		155	93	**248**

Research survey, 2014

4.3.2 Administration of the pilot study test

A time table for administering the test to the various schools was agreed with the respective heads of Biology department. Prior each administration of the test the purpose and importance of the test was explained to students. In each school, the administration was done simultaneously in the classes under the supervision of Biology teachers and a researcher within school time. Students were encouraged to attempt all 35 questions. Student scores were not reported back to teachers and were used for the purpose of item analysis, validity and reliability estimation only. All test papers were collected at the end of each test pilot and the teachers were not allowed to keep or make copies. Test scripts were scored by allocating a single mark for correct response and no mark for a wrong, omitted or a choice of more than one alternative per question. The total correct scores were determined and percentage of score out of the total number of possible scores (the total number of items) calculated. Both the raw scores and percentages for each subject were entered into a computer data analysis application for analysis.

4.3.3 Ethical issues

Before piloting the tool and the actual students assessment, permission or research clearance letter to carry out the study in Morogoro municipality was sought from the University of Wuppertal. The letter was sent by the researcher to the Regional Administrative Officer of Morogoro and then to the District Executive Directors (DED) of Morogoro municipality to grant permission of visiting the sampled schools. Ethical dimensions for individual participants were also considered. All the participants were duly informed of the objectives of the study before the test was administered. Participants were informed of their right to decline from participating in the study if they wish. The test scripts were handled by the researcher and the supervisor only. After marking, the scripts were stored in a safe place and they will be destroyed one year after completion of this study. The performance of each participant on the test was treated with high confidentiality. Pilot testing took one month and was strictly voluntary. The identity of all students participating remained confidential to ensure they remain anonymous to external populations. The only demographics collected were gender and grade levels. All participants will have the right to access their test results if they wish upon request.

4.3.4 Data analysis procedures

Data gathered in the pilot study were initially subjected to screening for normality and determining their suitability for parametric analyses. In determining the levels of science process skills acquisition amongst the students in terms of overall integrated science process skills and in each of the specific science process skills descriptive statistics were employed. This means the mean of scores and standard deviations were calculated from the SPSS. Difficulty index (P) was also calculated through SPSS package as the percentage of students who selected the correct response in each item. Discrimination index of each item was calculated using a scientific calculator as the difference of difficulty index of the 27% (164 students) higher achievers (who performed well on the test) and that of the 27% (164) lower achievers in each item. Indices of distracters were also calculated using a scientific calculator as the proportion of students who choose each answer option.

Pearson product moments (r) were also computed from the SPSS application package to indicate the construct and concurrent validities of the developed BPST instrument. The internal consistent correlation coefficient of reliability which indicate consistence (Cronbach Alpha (∝)) of the items in measuring the same ability or trait was also computed direct using computer statistical package SPSS. While the standard error of measurement (SEM) was calculated using the formula given below

SEM = SD⋎1-r

Where SEM= Standard error of measurement

SD = Standard deviation

r= reliability coefficient

Both the Grade level readability score and ease *readability* of BPST instrument were calculated. In both cases, the Flesch-Kincaid readability formulas were employed. The *following Flesch* steps and formula were followed when calculating the Grade Level (GL) readability score

Step 1: Calculating the average number of words used per sentence.

Step 2: Calculating the average number of syllables per word.

Step 3: Multiplying the average number of words by 0.39 and add it to the average number of syllables per word multiplied by 11.8.

Step 4: Subtracting 15.59 from the result.

The ***Flesch-Kincaid Grade Level*** *readability score* formula is:

FKRA = (0.39 x ASL) + (11.8 x ASW) – 15.59

The Flesch-Kincaid Ease Readability (RE) formula on the other hand was used to determine ease readability of the developed instrument. The specific mathematical formula is: **RE = 206.835 – (1.015 x ASL) – (84.6 x ASW)** ,

where

RE= Readability Ease

ASL = Average Sentence Length (i.e., the number of words divided by the number of sentences)

ASW = Average number of syllables per word (i.e., the number of syllables divided by the number of words) Flesch

4.4 RESULTS AND DISCUSSION ON DEVELOPMENT AND VALIDATION OF BIOLOGY PROCESS SKILLS TEST (BPST)

4.4.1 Introduction

This section presents and discusses the key findings of the study which sought to develop and validate an integrated science process skills test specific for advanced level students taking Biology. The procedures for test development as outlined in section 4.2. were used in test development and the statistical methods as outlined in the section 4.3.4 of the methodology were used to analyze the data collected from a pilot study. The section starts with the results of initial validation by the experts from the University of Wuppertal-Germany and Morogoro-Tanzania in its section 4.4.2 before discussing pilot findings. The pilot study findings are presented in section 4.4.3 in the following series; section 4.4.3.1 item response patterns, 4.4.3.2 difficult indices, 4.4.3.3 discrimination indices, 4.4.3.4 distracters analysis, 4.4.3.5 descriptive statistics, 4.4.3.6 reliability and standard error of measurement, 4.4.3.7 validity evidence of the test, and 4.4.3.8 readability of the test. Section 4.5 is the summary of the characteristics of the developed test as obtained from the pilot study and at the end of the chapter, section 4.6 is the discussion of the qualities of BPST and conclusion.

4.4.2 Results of initial validation of BPST by experts

4.4.2.1 First level validation of the test draft

The first level of test validation was the validation through experts' opinions (see the procedures in section 4.2.1 of the methodology). Six science educators (two (02) from the University of Wuppertal and four (04) from Kilakala secondary schools in Morogoro Tanzania) with experience in both test construction and the science process skills reviewed the test draft. Two reviewers from the University of Wuppertal were the first to provide their general comments with regard to test items. There was a general consensus with respect to the face validity of the test draft and content. The draft initially had 43 items while the aim was to prepare a test having 35 items. Hence, the test items were further discussed with four experienced teachers in the department of Biology at Kilakala secondary school. This discussion was done in order to qualify and disqualify some items in the draft and remain with the required 35. As shown in table 4.5 below, some of the items were found to be assessing students' conceptual understanding than the processes of science. Table 4.5 summarizes how the eight items that were removed from the test draft and the reasons provided by experts.

Table 4.5 Items that were removed in the initial validation process

Sn	Item requirement	Process skill	Reason for removal
I	The item required the student to make an interpretation about a young friend who and in the hospital, they discover his mitochondria can use only fatty acids and amino acids for respiration, and his cells produce more lactate than normal.	Data analysis and interpretation	It was seen as more cognitive than process skill
Ii	The item measured whether the student is able to use data from a real or simulated population(s), based on graphs or models of types of selection, to predict what will happen to the population in the future.	Data analysis and interpretation	Measures interactions of process skills
Iii	The item measured if a student is able to use data and refine observations and measurements regarding the effect of population interactions on patterns of species distribution and abundance	Identifying variables	It is more cognitive than process skill
Iv	The item required learners to define operationally the success of safety advertising to reduce accidents at schools. The hypothesis that safety advertising will reduce schools accidents where each school will use a different number of safety posters to see if the number of accidents is reduced and keep records	Defining operationally	Measures very low level process skill
V	The assessed if the student is able to design a plan for collecting data to support the scientific claim that the timing and coordination of physiological events involve regulation.	Designing of experiments	It is more cognitive than process skill
Vi	The item asked students the best way of getting sufficient amount of DNA from human skull extract unearthed by palaeontologists when a small fragment of the scalp tissue was still attached to it	Designing of experiments	It is more cognitive than process skill
Vii	This item needed student to choose the best hypothesis in a situation where cage with male mosquitoes in it has a small earphone placed on top, through which the sound of a female mosquito is played. The question was why all the males immediately fly to the earphone and thrust their abdomens through the fabric of the cage	Hypothesis formulation	It had weak distractors and It is more cognitive than process
Vii i	This was a respiration question which needed students make a hypothesis from the experimental evidence which shows that the process of glycolysis is present and virtually identical in organisms. Three domains of organisms , Archaea, Bacteria, and Eukarya were taken as a case study	Hypothesis formulation	It is more cognitive than process skill

Research survey data, (2014)

This panel discussion was an attempt to make sure that he developed instrument has strong validity and reliability levels. To simplify scoring and analysis of scores, the qualified items were arranged in clusters such that the first seven items measured hypothesis formulation skill, the second seven items measured the ability in identifying and controlling variables, and the third seven items were for the design of experiments. The fourth quarter of seven items (question number 21-28) measured ability in the interpretation of data and the last seven items measured students skills in defining operationally. Table 4.6 below summarizes the distribution of questions in the Biology process skills test and the specific process skill they measure (see Appendix I).

Table 4.6 Items in the BPST draft and skills they measure

	process skills	Questions in the test measuring it
I	Identifying and controlling variables	Question No; 8, 9, 10, 11, 12,13 and 14
Ii	Stating hypotheses	Question No; 1, 2, 3, 4, 5, 6 and 7
Iii	Operational definitions	Question No; 29, 30, 31, 32, 33, 34 and 35
Iv	Graphing and interpreting data	Question No; 22, 23, 24, 25, 26, 27 and 28
V	Experimental design	Question No; 15, 16, 17, 18, 19, 20 and 21

BPST, (2014)

4.2.2.2 Content validity of BPST instrument

Studies have highlighted the importance of involving experts in test content validation and judges who are not experts in the measure but are specialized in the construct of interest or knowledgeable of the discipline (Lynn, 1986 & Davis, 1992). Content validity of the developed instrument now having 35 items was determined by a panel of four (04) experienced Biology teachers who have worked extensively with secondary education students in Tanzania (see the procedures in section 4.2.1 of the methodology). The recommendation is to select at least three judges for each item (Lynn, 1986). The objective was to analyze and find out the extent to which the items created are representative of the target construct and the degree to which such items represent the facet of the construct they were developed for. A content validation form (Appendix III) was prepared as a guide for the 04 reviewers. The classic criteria established by Angleitner et al. (1986) were used as a reference. The items were assessed on the basis of the following criteria; accuracy or clarity (assessment of whether the item is properly understood and the extent to which the item is concise/accurate/direct), ambiguity (judgment on the chances that the item can be interpreted in different ways), and relevance or objectivity of individual items as they relate to an identified science process skills learning objectives. In addition to the guide, the raters were given the test items and a list of the test objectives, to check the validity of the test by matching the items with the corresponding objectives. The content validity of the instrument was obtained by determining the extent to which the raters agreed with the test developer on the assignment of the test items to the respective objectives (Dillashaw & Okey, 1980 & Burns et al. 1985). Items were scored on a scale of 0 (low) to 3 (high). Averages of reviewer scores for the 35 items are presented in table 4.7 below.

Table 4.7 Mean Average of Reviewers' Scores by Item (content validity)

Item #	Accuracy	Clarity	Objectivity	Item #	Accuracy	Clarity	Objectivity
1	2.4	2.6	3.0	**19**	2.2	3.0	2.8
2	2.8	3.0	2.5	**20**	3.0	3.0	3.0
3	2.5	3.0	3.0	**21**	2.6	2.5	3.0
4	2.6	2.5	2.5	**22**	2.5	2.4	3.0
5	3.0	2.5	3.0	**23**	2.8	3.0	2.5
6	2.6	2.6	3.0	**24**	2.5	3.0	3.0
7	2.0	2.5	2.5	**25**	3.0	2.5	2.5
8	2.5	2.4	3.0	**26**	2.0	2.5	2.5
9	2.8	3.0	2.5	**27**	2.5	2.4	3.0
10	2.5	3.0	3.0	**28**	2.8	3.0	2.5
11	2.6	2.5	2.5	**29**	2.5	3.0	3.0
12	3.0	2.5	2.5	**30**	2.0	2.4	3.0
13	2.2	2.6	3.0	**31**	2.6	2.6	2.8
14	2.0	2.5	2.5	**32**	2.8	3.0	2.5
15	2.5	2.4	3.0	**33**	2.5	3.0	3.0
16	2.8	3.0	2.5	**34**	3.0	2.4	2.6
17	2.5	3.0	3.0	**35**	2.5	3.0	3.0
18	3.0	2.5	2.5				
				Mean	**2.5**	**2.6**	**2.8**

Research survey data (2014)

As shown in table 4.7 above, for the criterion of accuracy, the 04 reviewers assigned fairly consistent ratings across the items. When reviewers were asked to evaluate each item on a scale of 0 (low) to 3 (high) in regards to how accurately the item described the objective. All four reviewers scored items at a 2 and a 3 level. The average score by an item of all 4 experts ranged from 2.2 to 3.0, with a mean score of 2.5. This is equivalent with the content validity of 83% or 0.83. Item clarity of the 35 items retained for inclusion on the final test was also fairly high. Of the 35 items, 29, or 83%, received an average score of 2.6 or more. Few items that received a rating of 2 were slightly modified. The mean score for the criterion of clarity of 35 items was 2.6 or 87%. As the test was devised to be used across multiple settings, the objectivity of the item was another important criterion of quality. Using the same 0 to 3 scale for accuracy and clarity, the experts scored objectivity very highly. All item average scores were 2.5 or higher. The mean average for objectivity across all items was 2.8 or 93%.

The overall mean or average of rater in all three criteria (83% accuracy, 87% clarity, and 93% objectivity) is 88% or 0.88. This concurrence of raters was taken as evidence of content validity and objectivity of scoring. Revisions were made on those items where modifications were suggested. Modifications included changes in wording and sentence length in order to provide additional explanation for specific terms, i.e., the manipulated variable or the condition which was changed, the responding variable or the condition which is the measured outcome of the experiment, and the controlled variable or the condition that was kept constant. The data obtained show that the process of obtaining content validity evidence leads to an improvement of the items created both regarding the formal wording aspects and the theoretical representativeness-relevance of such items. Thus, obtaining content validity evidence makes it possible from the outset to provide empirical data supporting the construction/adaptation process (Sireci, 1998), which also facilitates the subsequent stages. The responses of the four reviewers were consistent on almost all items in terms of indicating the correct answer and keying to a process skill objective.

4.4.3 Findings and discussion of the pilot study data

4.4.3.1 Item response patterns

After incorporating experts' opinions, the test version was subjected to pilot study with 610 Form IV, V, and VI students in the municipality of Morogoro. Section 4.3 of the methodology part explains sample size and schools involved in the pilot study while table 4.2 shows sample distribution of students based on their sex and grade levels. A pilot study aimed at generating data for determining the effectiveness of each item. Pilot data were also used to calculate validity and reliability of the test. The results from the pilot study were analyzed and item response patterns from students determined. The aim of determining item response patterns of the whole test was to determine how well students have distributed across item option (alternatives) and get an insight of whether there are some items or distracters need to be changed, modified or removed.

Table 4.8 below is a summary of item response patterns. The table (table 4.8) shows that only one question (question number 07) which measured students' ability in formulating and identifying hypothesis had a difficulty index outside the recommended range of 0.2 - 0.8 (Nitko, 1996). The item had a difficulty index of 0.18. The difficulty indices ranged from 18.00% (the difficult item, that is item 07 as we said) to 58.70% (the easiest item, that is item number 35). This question (question 7) was based on apparatus set up to investigate the effect of light intensity on the rate of photosynthesis in an aquatic plant. The candidates were required to select a hypothesis that could be tested using the apparatus shown in the diagram. Some candidates choose the hypothesis in the form of an aim or an expected observation. Another common error was selecting the option that stated *'light intensity affects the rate of photosynthesis'*. The correct option for the hypothesis was an option (D) *that an increase in light intensity results in an increase in the rate of photosynthesis*. This question was subsequently replaced in the final version of BPST with another question which also measure students' hypothesis formulation skill in the area of plant growth after discussion Biology teachers. The item was also removed by the fact that there another question in the area of photosynthesis, it was seen as a repetition.

Even though 50% is the ideal difficulty index, however, the study aimed at retaining all items having difficulty indices from 20% to 80% as shown in table 4.8 below. Table 4.8 also shows the distribution of students' choices in each item response option. From the table, it can be seen all of the incorrect options, or the distracters, actually acted as real distracting. They were selected by quite a good number of students. Each distracter was selected by a greater proportion of the lower group than of the upper group resulting into an excellent item response pattern as shown in the table 4.8 below

Table 4.8 **Item response pattern**

Item	Key	A	B	C	D	Decision
1	D	45	140	85	274	Retain
2	B	60	323	137	123	Retain
3	D	87	80	101	265	Retain
4	A	255	99	141	115	Retain
5	C	79	93	276	162	Retain
6	C	101	114	241	154	Retain
7	A	168	190	113	139	Don´t Retain
8	C	251	110	163	86	Retain
9	B	178	92	80	260	Retain
10	B	118	92	269	121	Retain
11	C	277	108	95	130	Retain
12	D	105	86	268	151	Retain
13	C	131	162	226	81	Retain
14	B	156	233	118	103	Retain
15	D	165	100	86	259	Retain
16	C	115	88	151	256	Retain
17	A	210	106	120	174	Retain
18	B	110	276	97	127	Retain
19	B	107	240	90	173	Retain
20	D	230	89	121	170	Retain
21	D	152	193	109	156	Retain
22	A	118	170	208	112	Retain
23	A	178	166	176	90	Retain
24	D	170	156	181	103	Retain
25	A	160	104	188	158	Retain
26	C	110	124	221	155	Retain
27	C	177	150	93	190	Retain
28	D	181	169	108	152	Retain
29	B	230	109	116	155	Retain
30	A	276	108	79	147	Retain
31	B	261	110	110	129	Retain
32	A	95	106	110	299	Retain
33	A	122	94	107	287	Retain
34	D	96	272	108	134	Retain
35	A	57	99	358	96	Retain

Source: Research survey (2014)

4.4.3.2 Difficulty indices of items in the BPST

One of the statistics explored during classical item analysis with pilot study data was item difficulty. As it has been stated in section 4.4.3.1, difficulty indices did not range widely for the 35 items, this probably might be as a result of the inclusion of 339 (55%) lower level students (Form Four) in the pilot. Although they are taking Biology, but these students have never been exposed more to process skills than the higher level students. The rationale for inclusion of Form Four students in this stage was to generate data that would be used in estimating the construct validity of the test (BPST). Table 4.9 is a summary of difficulty indices of items in the BPST, the number of correct and incorrect responses and the decision made in each item as a result of analysis.

Table 4.9 Difficulty index of each item in BPST

Item	Key	P	Number correct	Number incorrect	Incorrect percentage	Decision
1	D	0.45	274	336	55	Retain
2	B	0.53	323	287	47	Retain
3	D	0.43	265	345	57	Retain
4	A	0.44	269	341	56	Retain
5	C	0.59	359	251	41	Retain
6	C	0.43	263	347	57	Retain
7	A	0.18	113	497	82	Don´t Retain
8	C	0.46	279	331	54	Retain
9	B	0.53	320	290	47	Retain
10	B	0.54	333	277	46	Retain
11	C	0.63	385	225	37	Retain
12	D	0.58	352	258	42	Retain
13	C	0.54	330	280	46	Retain
14	B	0.75	455	155	25	Retain
15	D	0.29	174	436	71	Retain
16	C	0.45	277	333	55	Retain
17	A	0.58	355	255	42	Retain
18	B	0.68	415	195	32	Retain
19	B	0.66	403	207	34	Retain
20	D	0.40	243	367	60	Retain
21	D	0.36	220	390	64	Retain
22	A	0.34	208	402	66	Retain
23	A	0.27	166	444	73	Retain
24	D	0.25	156	454	75	Retain
25	A	0.30	188	421	70	Retain
26	C	0.25	155	455	75	Retain
27	C	0.24	150	460	76	Retain
28	D	0.29	181	429	71	Retain
29	B	0.37	230	380	63	Retain
30	A	0.45	276	334	55	Retain
31	B	0.42	261	349	58	Retain
32	A	0.49	299	311	51	Retain
33	A	0.47	287	323	53	Retain
34	D	0.44	272	338	56	Retain
35	A	0.58	358	252	42	Retain

Source: Research survey (2014)

As table 4.9 above show, item difficulty ranged from 18% item 07 (most difficult) to only 75% item number 14 (the easiest item). The mean of items´ difficulty indices was only 0.447. Nevertheless, this indicates that the test contains items of various difficulty levels and that students exhibited a broad range of skills levels. The average value of "P" of the test is 0.447 which reflects that the test items are capable to differentiate the level of scientific skills of the students to a greater extent. As it can be seen in table 4.9 above, the value of difficulty indices of items number 14, 18 and 19 can be treated as the easiest item. On the other hand, items no.7, 23, 24, 26, 27 and 28 are difficult items with respect to the sample of students. Table 4.9 further indicates that only one question, question number 07 had difficulty index outside the acceptable range and is was disqualified hence removed (see also section 4.3.3.1). Although it is quite possible for test takers to score in the upper ranges on criterion-referenced tests, Kehoe (1995) suggests that items answered correctly by 30% to 80% of test takers are good target difficulty ranges for discriminating knowledge. Allen & Yen (2001) on the otherhand suggested that, items with difficulty levels ranging from 0.20 to 0.90 to be retained in the final draft. Such items would have average difficulty and would be ideal (Allen & Yen, 2001). In this study however, items with difficulty levels above 0.80 were considered to be very easy and those below 0.20 were considered to be very difficult and are not preferred (see section 4.2.3).

Table 4.9 further shows a range of difficulty levels among the five science process skills clusters. The first cluster, hypothesis formulation contains items with a difficulty indices ranging from 0.18 (item 7) to 0.59 (item 5). The second cluster, demonstrating knowledge of identifying and controlling variables, ranges from 0.46 (item 08) to 0.75 (item 14), while cluster three, the design of experiments, had indices ranges from 0.29 (item 15) to 0.68 (item 18). Cluster four, demonstrating knowledge of analyzing and interpreting data on the other hand has items having difficult indices ranging from 0.24 (item 27) to 0.34 (item 22) and the last cluster (cluster number five), demonstrating the ability in defining variables operationally, ranges from 0.37 (item 29) to 0.58 (item 35). Calculation of the means (averages) of different science process skills clusters showed that sample students for pilot study did relatively better in items measuring ability in identifying and controlling variables, with the mean of indices of 0.57 and they did poorly in items measuring ability in data analysis and interpretation with the mean of indices of 0.28. Items measuring students' ability in hypothesis formulation had a mean of indices of 0.43 and those measuring experiment designs had the mean of 0.49, while defining operationally had a mean of 0.46.

4.4.3.3 Discrimination indices of BPST

The discrimination index, or point biserial correlation, compares performance on a given item from top scoring students with a performance from students in the bottom group. In this study, item discrimination indices were obtained by using the upper 27% (164 students) and lower 27% (164 students) of the sample. Point biserial correlation indices of item discrimination ranged from 0.26 item number 7 to 0.70 item number 35. The average index of discrimination was 0.48. The values of discriminatory power (D) shows that there is not a single item having zero or negative values. Item no. 7 is the only item having the value less than 0.30; (the acceptable range in this study) which demanded a major change or replacement. The item had the value of 0.26 and it was subsequently replaced. The item no. 17, 24, and 27 demanded a minor change in the item as they had weak indices of discrimination (see table 4.10 below).

Point biserial correlation indices based on the five specific science process skills clusters were also calculated. The findings show that the first cluster, hypothesis formulation had indices of discrimination ranges of 0.26 (item 7) to 0.59 (item 2). The second cluster, demonstrating knowledge of identifying and controlling variables, had indices ranging from 0.48 (item 9) to 0.67 (item 14), while cluster three, the design of experiments, had indices ranging from 0.35(item 17) to 0.62 (item 18). Cluster four, demonstrating knowledge of analyzing and interpreting data on the other hand contains items having discrimination index ranging from 0.34 (item 27) to 0.48 (item 22) and the last cluster (cluster number five), demonstrating the ability in defining variables operationally, had indices ranging from 0.53 (item 29)to 0.70 (item 35). Furthermore calculation of the means (averages) of different science process skills clusters showed that items measuring students´ ability in identifying and controlling variables had highest discriminating power with the mean of indices of 0.72, and items measuring skills in data analysis and interpretation had the lowest discrimination power with the mean of 0.39. Items measuring students' ability in hypothesis formulation, the design of experiments design and defining operationally had a mean of indices of discrimination of 0.46, 0.48 and 0.63 respectively.

Table 4.10 Discrimination indices of items in the BPST

Item	P	Number correct	Higher Achievers (27%)	Lower group (27%)	P (H)	P(L)	D index	Decision
1	0.45	274	108	36	0.65	0.21	0.44	Retain
2	0.53	323	130	34	0.79	0.20	0.59	Retain
3	0.43	265	101	32	0.61	0.19	0.42	Retain
4	0.44	269	119	40	0.72	0.24	0.48	Retain
5	0.59	359	136	53	0.83	0.32	0.51	Retain
6	0.43	263	124	34	0.76	0.20	0.56	Retain
7	0.18	113	56	13	0.34	0.08	0.26	Don´t Retain
8	0.46	279	125	43	0.76	0.26	0.50	Retain
9	0.53	320	130	41	0.79	0.31	0.48	Retain
10	0.54	333	138	41	0.84	0.25	0.59	Retain
11	0.63	385	143	43	0.87	0.26	0.61	Retain
12	0.58	352	137	36	0.83	0.21	0.62	Retain
13	0.54	330	132	37	0.80	0.22	0.58	Retain
14	0.75	455	149	37	0.90	0.23	0.67	Retain
15	0.29	174	88	21	0.54	0.13	0.41	Retain
16	0.45	277	109	26	0.66	0.16	0.50	Retain
17	0.58	355	104	45	0.63	0.27	0.35	Retain
18	0.68	415	143	42	0.87	0.25	0.62	Retain
19	0.66	403	129	48	0.79	0.29	0.50	Retain
20	0.40	243	119	33	0.73	0.20	0.53	Retain
21	0.36	220	95	23	0.58	0.14	0.44	Retain
22	0.34	208	103	25	0.63	0.15	0.48	Retain
23	0.27	166	82	20	0,50	0.12	038	Retain
24	0.25	156	77	18	0.47	0.11	0.36	Retain
25	0.30	188	93	22	0,57	0.13	0.43	Retain
26	0.25	155	77	18	0.47	0.11	0.36	Retain
27	0.24	150	74	18	0.45	0.11	0.34	Retain
28	0.29	181	89	21	0.54	0.13	0.41	Retain
29	0.37	230	114	27	0.70	0.16	0.53	Retain
30	0.45	276	136	33	0.83	0.20	0.63	Retain
31	0.42	261	129	31	0.79	0.19	0.60	Retain
32	0.49	299	148	35	0.90	0.21	0.69	Retain
33	0.47	287	142	34	0.87	0.21	0.66	Retain
34	0.44	272	134	32	0.82	0.20	0.62	Retain
35	0.58	358	152	37	0.93	0.23	0.70	Retain

Source: Research survey (2014)

On the basis of the results, it can be claimed that there is no a single item in BPST showing negative value with respect to the sample selected. Table 4.10 above indicates that only item number 7 failed to meet acceptable index range. A range of opinions exists among researchers as to whether it is advisable to discard items that appear to be poorly performing. In general, researchers who analyze measurement results from the affective domain suggest deleting low performing items. Some researchers (Kehoe, 1995) suggest items with a discrimination index below 0.15 are reviewed and either revised or withdrawn. Popham (2002) however, contends that items may be left in if they are well written and satisfies the objectives. Those developing cognitive, criterion-referenced tests argue that the researcher should consider the importance of the knowledge of the objective over the performance of the item. After thorough scrutiny, the researcher decided to delete the item as it was believed it did discriminate among knowledge levels.

5.3.4 Analysis of response options of BPST items (distractor analysis)

Analysis of distracters was also performed using pilot study data, and it aimed at assessing the plausibility of distractors. In a simple approach to distractor analysis, the key is equivalent to the item p-value, or difficulty. If the proportions are summed across all of an item's response options they will add up to 1.0, or 100% of the examinees' selections. In this study, four patterns as suggested by Hills (1981) guided researcher´s analysis of item distractors. These included (1) at least 5% of examinee should select every distractor, (2) the right answer should be selected much more frequently by the examinees in the upper group than those in the lower group, (3) each distractor must be chosen more by the lower-scoring examinees than the higher-scoring ones, and (4) it is desirable that the difficulty level of each item is similar to the optimal proportions.

Table 4.11 below summarizes the distribution of students in each item alternative in the Biology process skills test (BPST). The "percent choosing" columns provided the basis for distractors analysis. According to the table, all response options were plausible and were able to attract quite a number of students. As a result, all distractors (alternatives) qualified to be retained in the final draft. The incorrect answer options were chosen more by the lower-scoring examinees than the higher-scoring ones. Every alternative was chosen at least by 5% and a number of items demonstrated a good dispersal among choices with at least 15% choosing each alternative as shown in the Table 4.11 below.

Table 4.11 Distribution of students in each item alternative

Item	Key	A	B	C	D	%A	%B	%C	%D	Decision
1	D	139	146	98	274	23	24	16	37	All qualified
2	B	83	323	248	188	14	15	41	31	All qualified
3	D	120	108	130	265	20	18	21	41	All qualified
4	A	269	99	141	115	42	16	23	19	All qualified
5	C	79	93	359	162	13	15	45	27	All qualified
6	C	101	114	263	154	17	19	40	25	All qualified
7	A	113	190	113	139	28	31	19	23	All qualified
8	C	251	110	279	86	41	18	27	14	All qualified
9	B	178	320	80	260	29	15	13	43	All qualified
10	B	118	333	269	121	19	15	44	20	All qualified
11	C	277	108	385	130	45	18	16	21	All qualified
12	D	105	86	268	352	17	14	44	25	All qualified
13	C	131	162	226	81	21	27	37	13	All qualified
14	B	156	455	118	103	26	38	19	17	All qualified
15	D	165	100	86	174	27	16	14	42	All qualified
16	C	115	88	277	256	19	14	25	42	All qualified
17	A	355	106	120	174	34	17	20	29	All qualified
18	B	110	415	97	127	18	45	16	21	All qualified
19	B	107	403	90	173	18	39	15	28	All qualified
20	D	230	89	121	243	38	15	20	28	All qualified
21	D	152	193	109	220	25	32	18	26	All qualified
22	A	208	170	188	112	19	28	34	18	All qualified
23	A	166	178	176	90	29	27	29	15	All qualified
24	D	170	156	181	156	28	26	30	17	All qualified
25	A	188	104	160	158	26	17	31	26	All qualified
26	C	110	124	155	230	18	20	36	25	All qualified
27	C	177	93	150	190	29	25	15	31	All qualified
28	D	152	169	108	181	30	28	18	25	All qualified
29	B	109	230	116	155	38	18	19	25	All qualified
30	A	276	108	79	147	45	18	13	24	All qualified
31	B	110	261	110	129	43	18	18	21	All qualified
32	A	299	106	110	95	16	17	18	49	All qualified
33	A	287	94	107	122	20	15	18	47	All qualified
34	D	96	134	108	272	16	45	18	22	All qualified
35	A	358	99	57	96	9	16	59	16	All qualified

Source: Research survey (2014)

It can be concluded that the distractors in the Biology process skills test (BPST) meets criteria as set by Hills (1981). They were selected by at least 5% of examinee and that the right answers (key) were selected much more frequently by the examinees in the upper group than those in the lower group. The proportion of the examinees who select each of the distracters can be very informative. For example, it can reveal an item mis-key. Messick (1995) recommend that whenever the proportion of examinees who selected a distractor is greater than the proportion of examinees who selected the key, the item should be examined to determine if it has been mis-keyed or double-keyed.

A distractor analysis can also reveal an implausible distractor. If examinees consistently fail to select a given distractor, this may be evidence that the distractor is implausible or simply too easy (Hills, 1981).

4.4.3.5 Descriptive statistics of BSPT instrument

For additional SPSS analyses, the entire Biology process skills test down into its five subscales of process skill objectives. The mean scores end standard deviations on the BPST total and each subscale and overall students were calculated and summarized in Table 4.12 below. The mean and standard deviation for all pupils on the 35 items were 16.2 and 5.41 respectively. The score ranged from 6 (the lowest) to 28 (the highest). For overall students, correct response percentages were highest for the process skills of identifying and controlling variables (49.2.0%) and were lowest for the process skills of analysis and data interpretation (32%). In other words, learners from the different grade levels found the skill of identifying and controlling variables easier than other skills (Table 4.12). While fewer learners from the different performance categories in each grade selected the correct option for items related to the skills of analyzing data and designing experiments (Table 4.12). This finding is suggesting the possibility of learners having less experience in designing experiments, and the likelihood of the use of prescribed experimental designs in science classes. The mean of a raw score of the subtest was low, indicating that the students found the subtest fairly difficult. These findings were expected because the involvement of 339 (55%) Form four students in the pilot study. Due to their lower education level, Form IV students are presumed to have a low level of higher order science process skills than their counterpart Form V and VI. The positively skewed scoring distribution as shown in Table 4.12 also indicated that most of the students taking the subtest obtained low scores. Skewness is the extent to which a distribution of values deviates from symmetry around the mean. A value of a positive 0.1153 skewness means that there were a greater number of smaller values than mean (Cramer, 1997). Table 4.12 below is a summary of descriptive statists and the difficulty levels of each process skill objective as a subscale of the BPST.

Table 4.12 Descriptive statistic of BPST and its component skills

	Formulating hypothes	controlling variables	experiments design	interretation of data	operational definition	Total test
N	610	610	610	610	610	610
Range	7	7	7	7	7	7
Minimum	0	0	0	0	0	6
Maximum	7	7	7	7	7	7
Mean	3.40	3.45	3.35	2.24	3.42	16.2
Standard deviation	1.53	1.62	1.47	1.66	1.73	5.4
Skewness	-0.025	0.099	0.46	0.22	0.099	0.153
Kurtosis	-0.593	-0.833	-0.55	-0.722	-0.769	-0.898

Source: Research survey (2014)

The descriptive statistics obtained by field tests indicated the Mean performance on the BPST test was comparatively low (16.2 out of 35). This is equivalent to 46% percentagewise. The standard deviation of 5.75 and a range of 07 to 28 indicated that the test differentiated students of differing ability. This observation was supported by the indices of reliability, item difficulty, and item discrimination which provided measures of the suitability of the test scores for differentiating student performance. Nevertheless, these descriptive statistics summarized in the table 4.12 above indicates that the test contains items of various difficulty levels and that students exhibited a broad range of skills levels.

4.4.4 Reliability of the BPST instrument

4.4.4.1 Internal consistency reliability

The internal consistent reliability which indicates the consistency of items in measuring the same ability or trait was calculated using statistical package for social science SPSS version 16. Internal consistency reliability Cronbach value (∝) of 0.799 (approximated to 0.80) was obtained. Cronbach's alpha is the average value of the reliability coefficients one would obtain for all possible combinations of items when split into two half-tests. The coefficient of 0.80 obtained is well above the lower limit of the acceptable range of values for reliability [0.70 – 1.0] (Adkins, 1974; Hinkle et al. 2003), and it is within the range of reliability coefficients obtained from similar studies, such as; integrated process skill by Mungandi (2005) who obtained a reliability of 0.81, Dillashaw &

Okey (1980) who obtained a reliability of 0.89, Mattheis & Nakayama (1988) who obtained a reliability of 0.75, Onwu and Mozube (1992) who obtained a reliability of 0.84, process skills test by McKenzie and Padilla (1986) who obtained reliability of 0.84 and the recent one by Shahali & Halim (2010) who obtained a reliability of 0.808. A Cronbach alpha table 4.13 below as proposed by George & Mallery (2003) below provides the best way to interpret the meaning of Cronbach alpha values.

Table 4.13 A cronbach alpha table as proposed by George & Mallery (2003)

Cronbach's alpha	Internal consistency
$\alpha \geq 0.9$	Excellent
$0.9 > \alpha \geq 0.8$	Good
$0.8 > \alpha \geq 0.7$	Acceptable
$0.7 > \alpha \geq 0.6$	Questionable
$0.6 > \alpha \geq 0.5$	Poor
$0.5 > \alpha$	Unacceptable

Source: George & Mallery (2003)

The interpretation table shows that the closer Cronbach's alpha coefficient is to 1.0 the greater the internal consistency of the items in the scale. The reliability of BPST of 0.80 Cronbach value falls under a good range of internal consistency reliability level. However, as it has been noted by Cortina (1993) α value is also influenced by factors other than the test's operating characteristics. These factors warrant caution when interpreting the α statistic. According to Cortina (1993) the value α will be lower when the test is administered to a broadly heterogeneous population, and will be higher when evaluated using a more homogenous sample; b) α is affected by the length of the scale, so that longer tests may achieve satisfactory internal consistency despite low inter-item correlations; c) where α is extremely high, it may indicate redundancy among test items, suggesting the need to shorten the measure for the sake of parsimony and eloquence. Hence caution should be taken in interpreting because the large the value of α may not necessarily mean higher inter-item correlations.

Reliabilities of BPST by subtests were also computed from SPSS. The result shows that all 5 subtests of BPST had reliability (α value) of above 4. The subtest of defining operationally had the highest α value of 0.627, while the subtest of designing experiments had the lowest α value of 0.495. Formulation of hypothesis on the other hand α value of 0.510, controlling of variables had α value of 0.566 while interpreting data had α value 0.568 as summarized in the table 4.14 below.

Table 4.14 Reliabilities of specific BPST subtests

Subtest	Number of items	Reliability
Formulating hypothesis	7	0.510
Controlling variables	7	0.566
Design experiments	7	0.495
Interpreting data	7	0.568
Defining operationally	7	0.627
Total	**35**	**0.799**

The results from the table 4.14 above indicate that the performance of alpha is largely attributable to the reliabilities of the items that comprise a scale. It should however, be noted that while a high value for Cronbach's alpha indicates a good internal consistency of the items in the scale, it does not mean that the scale is unidimensional (Streiner, 2003). Factor analysis is a method to determine the dimensionality of a scale but is beyond the scope of this thesis.

4.4.4.2. Standard error of measurement (SEM) of BPST

Another way to express reliability is in terms of the standard error of measurement. In test theory, the standard error of measurement is the standard deviation of observed test scores for a given true score. This measure provides an estimate of how much an individual's score would be expected to change on re-testing with the same or an equivalent form of the test. The formula discussed in section 4.3.4 was used to determine the Standard Error of Measurements (SEM), to further estimate the reliability of the instrument. Standard error of measurement (SEM) on the other hand was calculated using the formula given below

SEM = SD⋎1-r

Where SEM= Standard error of measurement

SD = Standard deviation; in this case 18.95

r= reliability coefficient; in this case = 0.799

SEM= 18.95√1 – 0.799
= 16.12*0.448
= 7.22

This value (7.22) is relatively small, which means that the learners' obtained scores did not deviate much from their true scores. The smaller the standard error of measurement, the more reliable the results are (Nitko, 1996).

4.4.5 Validity of the BPST instrument

4.4.5.1 Concurrent (Criterion-related) validity of BPST

Another purpose of this study was to examine the concurrent validity of a new test instrument the Biology process skills test (BPST). Concurrent validity procedures are used to determine how well the test compares to another measure of the ability of the construct being assessed. In this study, a valid and reliable test of integrated science process skill (TIPS-II) by Burns et al. (1985) was used for comparison purpose. BPST and TIPS-II were concurrently administered to among 248 advanced level students in Kilakala and Alfa Gems secondary schools. The schools were purposefully involved because of having many advanced level Biology students as summarized in table 4.3 of the methodology. Concurrent validity was obtained by correlating the learners' scores obtained from the developed test (BPST) and those from TIPS-II. A correlational coefficient (r) between BPST and TIPS-II obtained through SPSS was 0.50. However, this value is below the acceptable range of value for a coefficient of correlation ($0 \geq 0.7$). This might be that, although TIPS II a valid test but is not knowledge domain specific test as BPST. TIPS II consists of Physics Chemistry and Biology items all together while BPST consists of questions related to Biology only. This correlation was nevertheless necessary to show that local learners performed differently on the two tests.

A multiple bivariate analysis using computer SPSS was conducted to determine which of the five selected clusters of science process skills (formulating hypotheses, defining operationally, identifying and controlling variables, design experiments as well as and interpreting data) from the BPST would have the highest correlation with the TIPS-II. Again the cluster of defining operationally had the highest correlation ($r = .62$) and the cluster of data analysis and interpretation had the lowest correlation 0.31. The researcher is in view that determination of this coefficient involving the use of the TIPS-II was not suitable for use in this specific case.

4.4.5.2 Construct validity of BPST instrument

Construct Validity refers to the ability of a measurement tool to actually measure the psychological concept being studied. Discriminant validity was calculated to provide the evidence of the construct validity of BPST. Discriminant validity, by logic, consists of providing evidence that two tests that do not measure closely related skills or types of knowledge do not correlate strongly (i.e a dissimilar ranking of students). Although there is no standard value for discriminant validity, according to Shavelson et al. (1999), a result less than 0.85 tells us that discriminant validity likely exists between the two scales. A result greater than 0.85, however, tells us that the two constructs overlap greatly and they are likely measuring the same thing. Therefore, we cannot claim discriminant validity between them. Construct validity compares two groups, one that has the construct and one that does not have the construct in question. If the group with the construct performs better than the group without the construct, that result is said to provide evidence of the construct validity of the test (Reckase, 1998).

In this study, construct validity of BPST was calculated by comparing the performance of Form IV to that of advanced level students (Form V and VI)assuming that Form V and VI students are more knowledgeable in science process skills than their counterpart Form IV. As shown in table 3.3 of methodology, the study involved 339 Form VI and 271 Form VI students. The 339 Form four come from kihonda, Educare, Mororgoro and Kigurunyembe schools while 271 come from Kilakala, Alfa Germs and Bigwa sisters. The mean score of Form IV was 13.4 while that of advanced level students was 19.3. SPSS bivariate analysis provided a correlation coefficient of 0.34 as discriminant validity between two groups of students. This dissimilarity in students' performance as proved from their correlation in the BPST instrument was taken as a proof that it has strong construct validity of BPST.

4.4.6 Readability issues of the developed BPST instrument

As it has already been described in sections, both the grade level readability score and reading age readability of BPST instrument were calculated. In both cases, the Flesch-Kincaid readability formulas were employed. The results of the two tests correlate approximately inversely; a text with a comparatively high score on the reading ease test should have a lower score on the grade level test. The *Flesch Reading Ease* score is part of the best-known readability scores.

4.4.6.1 Readability level of the developed instrument

The readability level score of the Biology process skill instrument was obtained using the Flesch reading ease (FKRS) formula as outlined in section 4.3.4. The computation of readability level was based on 15 item randomly sampled such that each process skill (hypothesis formulation, experiments design, interpreting data, controlling variables and defining operationally) provides three items. Words and sentences associated with graphs, charts, and tables were excluded from the texts that were used in the calculation of the index. The Flesch reading ease scale is rated from 0 to 100. A high readability value implies an easy to read text. The average sentence length (ASL) was determined and found to be 16.15 while the average number of syllables per word (ASW) was found to be 1.40. The data used and the calculation of the readability level of BPST instrument is shown below.

The average sentence length (ASL) = 16.15 and

The average number of syllables per word (ASW) = 1.40

Readability score = (206.835 - (1.015*ASL) - (84.6*ASW)

= (206.835 - (1.015*16.15) - (84.6*1.40) = 72.0 (approx)

The suggested range for a fairly easy readability level is 70 to 80 (Mungandi, 2005). The readability level of the developed instrument was found to be 72.0 (as seen in the calculation above). This readability level is on the higher end of the 'fairly easy readability range,' on Flesch's reading ease scale (see the Score mapping table section 4.2.3.9). Therefore, the readability level of the Biology process skills test instrument may be considered fairly easy.

4.4.6.2 Reading grade level of the developed instrument

The *Flesch-Kincaid Reading Age (FKRA)* formula on the other hand, was used to determine the suitable reading age suitable for the developed instrument (see section 4.3.4). The average sentence length (ASL) was determined and found to be 16.1529 while the average number of syllables per word (ASW) was found to be 1.4024. The data used and the calculation of the readability level of BPST instrument is shown below.

The average sentence length (ASL) = 16.1529 and

The average number of syllables per word (ASW) = 1.4024

Reading Age (FKRA) is given by the following formula

FKRA = (0.39 x ASL) + (11.8 x ASW) - 15.59

Grade level score = (0.39*ASL) + (11.8*ASW) – 15.59

=(0.39*16.1529) + (11.8*1.4024) – 15.59
=7.3 approximated to 7

The results obtained from the calculation of the reading grade level for the developed instrument showed that the suitable reading level of the developed instrument is grade 7. It has to be noted that, BPST tool is meant for senior students (grade12 and 13) and not lower grades like grade 7. However, it is pertinent to point out that, the determination of the Flesh-Kincaid formula was based on grade levels from schools in American countries, where English is used as a first language. For most Tanzania learners, English is used as a second or third language. The actual grade level for Tanzania users of the test instrument is therefore likely to be higher than that suggested by the formula. This argument is supported by Stephens (2000) who states that at higher-grade levels, grade level scores are not reliable, because, background and content knowledge become more significant than style variables. They (grade levels) are therefore likely to under-estimate or over-estimate the suitability of the material. Nevertheless, in this study test readability level was determined to provide an estimation of the degree of which the learners would understand the text of the developed instrument, so that learners may not find the test to be too difficult due to language constraints. A reading level of 60 – 70 obtained above (section 4.4.6.1) was considered to be easy enough for the learners to understand the text of the test instrument.

4.4.7 Discussion on the qualities of the Biology process skills test (BPST)

The objective of this project was to develop a valid and reliable multiple-choice test of five integrated science process skills for students in grades 11 to 13 (ages 15 to 17 years). Test usefulness can be determined by considering its reliability, validity, discrimination power, authenticity, and practicality (Bachman, 2000). The findings support using BPST to assess higher school students' scientific literacy. Through careful attention to the standards for developing validity arguments of a psychometric test (Spector, 1992; Burton & Mazerolle, 2011), the study has provided comparative validity evidence related to test content, response process, and internal structure. The results of the iterative process of item construction, administration, and revision provide support that the BPST with the underlying conceptualization of scientific literacy that sought to be assessed. In addition, the development of this measure was guided by experts from the University of Wuppertal and the Ministry of Education in Tanzania with the sound conceptualization of scientific literacy based on the extant literature.

The BPST is intended to assess higher schools students' sense of field/discipline general scientific literacy (the processes of science). The BPST is designed to be administered in one class period of 55–60 minutes via a paper and pencil format. The descriptive statistics obtained by field tests indicated the mean performance on the BPST test was comparatively low. The mean and standard deviation for all pupils on the 35 items were 16.1 and 5.41 respectively. The score ranged from 6 (the lowest) to 28 (the highest). However, the standard deviation of 5.41 and a range of 22 indicated that the test differentiated students of differing ability.This observation was supported by the indices of reliability, item difficulty, and item discrimination which provided measures of the suitability of the test scores for differentiating student performance.

However, psychometric qualities of the developed BPST to some extent are in congruence with the qualities of many science process skills tests published. For example, it resembles the test of the integrated science process skills for secondary science students (TIPS) developed and published by Dillashaw & Okey (1980). These authors developed a valid and reliable test of integrated science process skills for students in secondary schools with an internal consistency the reliability of 0.89 across ability levels, socioeconomic levels, gender, and race. The mean of the second version of the TIPS was 18.99 and the standard deviation of 7.60. I addition the mean item discrimination index is 0.40 and the average item difficulty index is 53%. Each of these three test characteristics is within acceptable limits for reliable tests (Payne, 1974). An estimate of the readability index is 9.2. Although this value of readability level seems a bit high, it results from the necessary use of multiple syllable words associated with investigating.

The psychometric qualities of BPST also relate to the qualities of the test of integrated science process skills (TIPS-II) which was developed by Burns et al. (1985). This instrument (TIPS-II) is a multiple-choice items test which was generated for five objectives and was meant for middle and high school students. Results yielded a mean score of 19.14 and a total test reliability of 0.86. Mean difficulty and discrimination indices were 0.53 and 0.35, respectively. Split-test correlations coefficients between TIPS II and the original TIPS items were 0.86 and 0.90. TIPS II provided another reliable instrument for measuring process skill achievement.

As it has already been stated, reliability is an important component of a good psychological test and it looks consistency or reproducibility of test scores. It is the degree to which test scores for a group of test takers are consistent over repeated applications of a measurement procedure and hence is inferred to be dependable and repeatable for an individual test taker. As it can be seen in table 5.11 below, this project produced a test with an internal consistency reliability of 0.80 Cronbach alpha. The result on the reliability of the test (BPST) in this study resembles the quality of the test developed by Onwu & Mozube (1992). In their study which used Nigerian secondary education curriculum Onwu & Mozube (1992) developed and validated a science process skills test for secondary school science students with 36 multiple choice questions. Their test intended to measure students' skills in identifying and controlling variables, defining variables operationally, formulating hypotheses, interpreting data and in designing experiments (Onwu & Mozube, 1992). The test was considered valid in the Nigerian education context and had a reliability of 0.84. In another study which produced reliability coefficient almost similar to BPST, Shahali, and Halim (2010) developed a test of integrated science process (TISP) which consisted of 30 multiple-choice items specific to the science content defined in the Malaysian primary school science curriculum. The test assesses performance on a set of integrated science processes associated with planning investigations such as formulating hypotheses, operationally defining of variables, identifying and controlling variables as well as interpreting data. Total test reliability using Cronbach Alpha was measured at 0.808 (Shahali & Halim, 2010). The reliability quality of BPST also resembles a recent test of integrated process skills was developed by Mungandi of the joint center for Science Mathematics and Technology Education at the University of Pretoria, South Africa in 2005. The test is a thirty multiple-choice and its reliability was estimated using the split-half method with 1043 learners of grade 9, 10, and 11 to be 0.81. The reliability coefficient of the developed BPST is well above the lower limit of the acceptable range of values for reliability, and it is within the range of reliability coefficients obtained from similar process skills tests, such as that by Dillashaw & Okey (1980) who obtained a reliability of 0.89 and that by Burns et al. (1985) who also obtained a reliability of 0.84.

As it is for validity, different researchers also hold different views in regard to the acceptable ranges of reliability. According to Nunnally & Bernstein (1994), the acceptable reliability estimates of social sciences research instruments ranges from 0.70 to 0.80. However, research in the physical sciences typically demands more rigorous reliability standards, as the constructs involved are more concrete and easily defined. In both settings, acceptable reliability estimates should be congruent with the implications of the test scores. That is, higher stakes testing should have higher standards of instrument reliability (Nunnally & Bernstein, 1994).

Table 4.15 Summary of the characteristics of the developed BPST instrument

Test characteristic	Overall	Acceptable values
Difficulty index	0.447	0.2-0.8
Discrimination index	0.48	≥0.3
Content validity	0.88 (88%)	
Concurrent validity	0.51	≥0.7
Construct validity(Discriminant correlation coefficient)	0.34	≤0.85
Reliability(internal consistency)	0.80	≥0.7
Standard Error of Measurement(SEM)	7.22	Not specified
Readability level	72.0	60 - 70
Reading grade level	Grade 7	Grades 12-13

Research data (2014)

Validity is arguably the most important criteria for the quality of a test. It refers to the degree in which the test or other measuring device is truly measuring what it is intended it to measure. Concurrent validity was obtained by correlating the learners' scores obtained from the developed test (BPST) and those from TIPS-II. A Pearson's product-moment correlational analysis between BPST and TIPS-II obtained through SPSS computation was 0.50. However, this value is below the acceptable range of value for a coefficient of correlation (0≥0.7). This might be that, although TIPS-II a valid test but is not knowledge domain specific test as BPST. This correlation was nevertheless necessary to show that local learners performed differently on the two tests. A multiple bivariate analysis using computer SPSS was conducted to determine which of the five selected clusters of science process skills (formulating hypotheses, defining operationally, identifying and controlling variables, design experiments as well as and interpreting data) from the BPST would have the highest correlation with the TIPS II. Again the cluster of defining operationally had the highest correlation (r = .62) and the cluster of data analysis and interpretation had the lowest correlation 0.31

Discriminant validity of BPST, on the other hand, was calculated by comparing the performance of Form IV to that of advanced level students (Form V and VI)assuming that Form V and VI students are more knowledgeable in science process skills than their counterpart Form IV. As shown in table 3.3 of methodology, the study involved 339 junior students and 271 senior students.

The mean score of Form IV was 13.4 while that of advanced level students was 19.3. SPSS bivariate analysis provided a correlation coefficient of 0.34 as discriminant validity between two groups of students. This dissimilarity in students' performance as proved from their correlation in the BPST instrument was taken as a proof that it has strong construct validity of BPST. Discriminant validity, by logic, consists of providing evidence that two groups of students are not similar. Although there is no standard value for discriminant validity, according to Shavelson et al. (1999), a result less than 0.85 tells us that discriminant validity likely exists between the two scales. A result greater than 0.85, however, tells us that the two constructs overlap greatly and they are likely measuring the same thing. Therefore, we cannot claim discriminant validity between them. Construct validity compares two groups, one that has the construct and one that does not have the construct in question. If the group with the construct performs better than the group without the construct, that result is said to provide evidence of the construct validity of the test (Reckase, 1998).

The average value of difficult indices "P" of BPST is 0.447 which reflects that the test items are capable of differentiating the level of scientific skills of the students to a greater extent. Although it is quite possible for test takers to score in the upper ranges on criterion-referenced tests, Kehoe (1995) suggests that items answered correctly by 30% to 80% of test takers are good target difficulty ranges for discriminating knowledge while Allen & Yen (2001) suggests that items with difficulty levels ranging from 0.20 to 0.90 to be retained in the final draft. Such items would have average difficulty and would be ideal. In this study, items with difficulty levels above 0.80 were considered to be very easy and those below 0.20 were considered to be very difficult and are not preferred. The average index of discrimination of BPST items was 0.48. Some researchers (Kehoe, 1995) suggest items with a discrimination index below 0.15 are reviewed and either revised or withdrawn. Popham (2002), however, contends that items may be left in if they are well written and satisfies the objectives. Those developing cognitive, criterion-referenced tests argue that the researcher should consider the importance of the knowledge of the objective over the performance of the item.

4.4.8 Implications and rationale of having BPST

The science educational reforms of the 1960's and 1970's prompted the need to develop various test instruments for assessing both teachers and students' level of science process skills (Dillashaw & Okey, 1980; Mungandi, 2005). Since then, scientific process skills evaluations performed through paper and pencil multiple-choice test have been very famous (Burns et al. 1985; & Dillashaw & Okey, 1980). However, there has been a considerable reaction to and criticism towards traditional paper and pencil tests type of assessment especially in assessing the performance of tasks. These critics proposed that students can authentically be assessed in their ability to perform a task, like the science process skills when they perform the real task. According to Wiggins (1989), authentic assessment means that tests should involve real-life tasks, performances, or challenges that replicate the problems faced by a scientist, historian, or expert in a particular field; thus, they are complex tasks rather than drills, worksheets, or isolated questions. Harlen (1999) for example asserted that science process skills were inseparable in practice from the conceptual understanding involved in learning and applying science and played a central role in learning with understanding. Specifically, in the context of student performance in school science laboratory, Gronlund (1998) wrote that "if you want to determine whether students can conduct an experiment, let them conduct an experiment".

However authentic methods for assessing science process skills competence such as through laboratory practical work have a number of constraints particularly in the context of teaching and learning in large under-resourced science classes (Onwu & Stoffels, 2005 & Mungandi, 2005). This is particularly true of Tanzania secondary schools. On the other hand, science educationists have stressed the importance of integrating both the authentic and alternative methods in measuring outcomes of school science courses (Hofstein & Lunetta, 2004; Wiggins, 1987, & Wiggins, 1998). Wiggins (1998) for example asserted that even today in an era of highly emphasized standards approach to science education, assessment of students' performance in the science a teacher does not have to choose between authentic assessment and traditional objective tests, rather integrate . It is should be like that some mix of the two to best meet needs. Thus (Wiggins, 1998) recommended multiple and varied assessments be used so that i. a sufficient number of samples (topics) are obtained, and ii. a sufficient variety of measures is used. Hence this study developed the science process skills test that could be used both as an alternative measurement tool and an integral tool for authentic measurement of science process skills.

The results of the study show that the test characteristics of the developed instrument BPST fall within the acceptable range of values as summarized in table 4.15. Six science educators as reviewers of agreed about the objectivity of the test items as they were designed to measure and that the scoring key options as provided by the test developer were correct. This concurrence of raters was taken as evidence of content validity and objectivity of scoring. In addition to content validity, before the BPST test could be considered a highly valid measure of integrated science process skill ability, more validity data would be necessary, discriminant analysis for construct validity and correlation coefficient for concurrent criterion validity. This suggests that the developed instrument is valid and reliable enough to be used in measuring learners' competence in the stated science process skills. The advanced level secondary education Biology students are the target group. The paper and pencil group testing format does not require expensive resources, and it can easily be administered to large groups of learners at the same time, hence it may be concluded that the test is cost effective and convenient. The test could serve either as an alternate and equivalent process skills test or enlarge the pool of available items for process skills assessment in advanced level school grades. Since most of these indices fell well within the acceptable range for reliable tests, the BPST test can be used by teachers to obtain dependable student data as a basis for making decisions about individual students on item performance.

The search of available integrated science process skills tests showed the continued need and relative scarcity of a test geared to secondary school students and associated with any particular science curriculum. The development of the Biology Process Skills Test (BPST) is an attempt fills this void. The 35-item paper and pencil test can be administered to groups of students to obtain measures of process skill acquisition for use by classroom teachers; researchers, and evaluators. The test is specific to Biology curriculum although it may occasionally be used across the various disciplines of science. As well as assessing process skills competence of high school students, the test may also be a useful means of classroom-based research, evaluation of instruction and learning, and curriculum validation in the evaluation. The BPST instrument is a reliable instrument for diagnostic and/or summative assessment in science classes or research studies. The researcher recommends to teachers and evaluators to us this test in the following specific occasions:

i. Research studies in which the dependent variable is student acquisition of the science process skills.
ii. Teaching skills research where the effectiveness of certain teaching practices is measured by process skill achievement.
iii. Assessing process skill competency by classroom teachers where the process skills are an important outcome of science instruction. Tests of the effectiveness of materials or modules designed to aid students in learning science process skills.
iv. An alternative to a laboratory procedure as a way to assess process skill acquisition.

The BPST constructed seem to be a reasonably good measure of integrated SPS which is necessary for conducting scientific investigations. Findings from the analysis showed that improvement is still needed for five of the test items to ensure the instrument is reliable and useful. This instrument may not be limited in its usage as it is specific to the Tanzania school science curriculum.

4.4.9 Conclusion

The findings presented in this chapter support the use of BPST in assessing Biology students' scientific literacy level in science process skill. The BPST is designed to be administered in one class period of 55–60 minutes via a paper and pencil format. Through careful attention to the standards for developing validity arguments of a psychometric test (Spector, 1992 & Burton & Mazerolle, 2011), the study has provided comparative validity evidence related to testing content, response process, and internal structure. In addition, the development of this measure was guided by experts from the university of Wuppertal and from the Ministry of Education in Tanzania with the sound conceptualization of scientific literacy. The aim was to maximize discrimination, score variance degree of reliability, and evidence to support validity claims. The results of the iterative process of item construction, administration, and revision provide support that the BPST with the underlying conceptualization of scientific literacy that sought to be assessed.

Test usefulness is determined by considering its reliability, validity, discrimination power, authenticity, and practicality. The BPST items demonstrated good reliability and the items on each adhere to the recommended guidelines for percent correct, discrimination index, item-total correlation coefficients, and frequency distribution of distracters selected, all of which provide evidence for the strong internal structure of this measure. The test specifically measured students' performance on specific integrated science process skills which includes skills in i. formulating hypotheses, ii. defining variable operationally, iii. identifying and controlling variables, iv. planning investigations as well as v. analyzing and interpreting data. The test consists of 35 multiple-choice items and took into account realities of the Tanzania education system. It is a multiple choice, four-option format with an emphasis on the use of pictures and drawings to clarify and enhance items. The first version of the 35-item test of process skills in science was validated and administered to 610 eleventh-, twelfth-, and thirteenth-grade students to establish test reliability and to compute item difficulty and discrimination indices to aid in test revision. Evidence of content validity, construct validity, internal consistency reliability, difficult index, and discrimination index was investigated to prove its psychometric properties.

CHAPTER FIVE

STAGE TWO: ASSESSMENT OF THE LEVEL OF SCIENCE PROCESS SKILLS OF ADVANCED LEVEL BIOLOGY STUDENTS IN MOROGORO USING THE CONSTRUCTED BPST TEST

5.1 Introduction

In the second stage, the study employed the test that have been developed and validated in the first stage (BPST) to assess the level of Tanzania students in science process skills. Advanced level Biology students in Morogoro municipality were taken as a case study. This chapter explains the rationale, methodology , findings and discussion of the result obtained during the assessment students' level of science process skills using the BPST. As it has already been explained earlier that in 2005 Tanzania came up with the so called 'Competence Based Curriculum' which emphasized among other things, student's competence in science process skills. The curriculum emphasized the need of Tanzania science students to learn scientific subjects such as Biology, Physics and Chemistry in the same way science is done scientists. Despite such a dramatic shift in curriculum policy, little is known about whether or not the reform efforts are truly transforming the educational experiences of students. So this stage of the study was conducted in order to establish a base level of information on whether or not Tanzania students are acquiring the prescribed competences present in the competence based curriculum of 2005. The following research questions guided this phase

i. *What is the general performance of advanced level Biology students in the in science process skills test validated in the selected schools in Morogoro municipality?*

ii. *What is the performance of advanced level Biology students of Morogoro municipality in each of the five process skills (formulating hypotheses, defining operationally, identifying and controlling variables, design experiments as well as and interpreting data)and is there any statistically differences in the performance based on these specific skills*

iii. *Is there any statistically significant difference in the performance of advanced level Biology students in Morogoro municipality based on their demographic characteristics of sex and grade level?*

5.2 Methodology in this Stage

5.2.1 Research design

In assessing the level of science process skills level of students, descriptive research design was adopted. According to Krathwohl (1993), descriptive research design provides current information about conditions, situations, and events. Borg and Gall (1989) maintains that descriptive studies are used to find out "what is". A descriptive design was suitable at this stage because the study intends to provide descriptions of the level of science process skills of higher Biology students in Tanzania. Advanced Biology learners in the municipality of Morogoro were a representative case study. Descriptive statistics were also used to analyze overall test performance, and students performance by specific science process skills in an attempt to determine whether performance differs with the type of skill and by demographic variables of gender and grade levels. By so doing the researcher was able to produce information on what is going in the school system of Tanzania with regard to science process skills.

5.2.2 Sample size and sampling of participating schools

The population for this study was the advanced level Biology students (Form V and VI) who have Biology as one of their major subjects in secondary schools in the municipality of Morogoro. A list of advanced level secondary schools in the municipality of Morogoro and subjects they offer was provided by the district education officer for secondary education. According to the list, there are four secondary schools in different locations of Morogoro municipality which offers Biology for advanced level students. These schools are Kilakala, Alfa Germs, Bigwa Sisters and Lutheran junior seminary. These schools differ in terms of a number of students taking Biology. The subjects involved were all Form V and VI students who had undergone the revised science syllabus. Because of a need to assess a large sample of students, all students in these four schools were involved. It means that sample size, in this case, was equal to the population of senior Biology learners in the municipality of Morogoro. Furthermore, this implies that no any sampling technique was employed to obtain the appropriate sample size. The number of subjects in each grade level based on their gender has been shown in table 5.1 below. All subjects of this study had been given focused instruction on science process skills under the review science curriculum which required the systematic teaching of integrated science processes.

Table 5.1 Schools and number of students that participated in the study

Sex	School	Type of Student		Total
		Form VI	Form V	
Female	Kilakala	48	91	138
	Alfagerms	22	25	47
	Bigwa sisters	9	14	23
	Lutheran Junior Seminary	18	19	37
Total		**97**	**149**	**246**
Male	Alfagerms	25	37	62
	Lutheran Junior Seminary	20	25	45
Total		**45**	**62**	**107**
Grand Total		**142**	**211**	**353**

Research survey data (2014)

5.2.3 Data collection instrument

In assessing the knowledge level of integrated process skills of advanced Biology students in Morogoro, a Biology process skills test (BPST) developed and validated in the first stage of this study was used (see Appendix I). The test measures five (05) individual integrated scientific skills (identifying variables, stating hypotheses, operationally defining, designing investigations and analyzing and interpreting data) to advanced secondary school learners. The reliability of the instrument was established by the researcher in the year 2014 using 610 learners to be 0.80 (Cronbach's alpha). Concurrent validity of BPST was established by comparing students score in the process skills test (TIPS-II) by and Burns et al. (1985). The test has reliability coefficient well above the lower limit of the acceptable range of values for reliability, and it is within the range of reliability coefficients obtained from similar studies, such as those by Dillashaw and Okey (1980) who obtained a reliability of 0.89 and Burns et al. (1985) who also obtained a reliability of 0.84. Biology process skills test (BPST) has a readability index of 72.0. This high readability value implies an easy to read text to students who English is not their first language like Tanzania students.

The test fits with the context of Tanzania and the competence-based curriculum being implemented. Table 5.2 below shows the distribution of questions in the Biology process skills test and the specific process skill they measure.

Table 5.2 Items in the validated BPST and skills they measure

	process skills	**Questions in the test measuring it**
i	Identifying and controlling variables	Question No; 8, 9, 10, 11, 12,13 and 14
ii	Stating hypotheses	Question No; 1, 2, 3, 4, 5, 6 and 7
iii	Operational definitions	Question No; 29, 30, 31, 32, 33, 34 and 35
iv	Graphing and interpreting data	Question No; 22, 23, 24, 25, 26, 27 and 28
v	Experimental design	Question No; 15, 16, 17, 18, 19, 20 and 21

Additional demographic data of sex and grade levels were collected together with student responses in the Biology process skill test in the answer grid where students were supposed to write their sex and grade level.

5.2.4 Procedures and administration

Prior each administration of the test the purpose and importance of the test was explained to students. In each school, the administration of the test was done simultaneously in the classes under the supervision of Biology teachers and a researcher within school time. The test duration was one hour and was again voluntary. The identity of all students participating remained confidential. The only demographic information collected were sex and grade levels. Student scores were not reported back to teachers and were used only for the research purpose only. However, students were informed about their right to know their score. All test papers were collected at the end of each test and the teachers were not allowed to keep or make copies. Test scripts were scored by allocating a single mark for a correct response and no mark for a wrong, omitted or a choice of more than one alternative per question. The total correct scores were determined and percentage of score out of the total number of possible scores (the total number of items) calculated.

5.2.5 Grading system of students performance in BPST

The grading system for the advanced level students of Tanzania was adopted in grading Morogoro biology student scores in the process skills test. This scale has been upgraded by the National Examination Council of Tanzania (NECTA) in 2014 and classifies student scores into seven (07) classes. Grade A which ranges from 75% to 100% implies a very satisfactory or excellent performance, while B+ ranges from 60% t0 74% and implies satisfactory or good. The scale award to a student grade B who will score between 50% to 59% implying ´´Good or above average´´, and grade C (average), for a student who will score between 40% to 49%. Grade D stands for ´´Below average´´ or unsatisfactory performance and is awarded for a score between 30 – 39%. Grades E and F stands for poor and very poor respectively and are awarded to those students who would score between 20% - 29% and between 0%- 19% respectively. After marking, student scores were converted into percentages and classified into seven categories using the above criteria and presented in the format as shown in Table 5.3 below.

Table 5.3 Test scores grading system

Range of scores	Corresponding %	Grade	Description of the level of process skill
0- 6	0-19	F	Very unsatisfactory
7 - 10	20-29	E	Unsatisfactory
11- 13	30-39	D	Below average
14- 17	40 – 49	C	Average
18 - 20	50- 59	B	Satisfactory
21- 26	60-74	B+	Very Satisfactory (Very Good)
27 - 35	75-100	A	Excellent

Source: URT (2014)

5.2.6 Data analysis plan

Descriptive statistics were used to analyze overall test performance and students' performance by specific science process skills in an attempt to determine whether performance differs with the type of skill. Students' score were analyzed using SPSS version 21.0. Descriptive analysis of frequencies, percentages, means and standard deviations was used to categorize, organize and analyze student score from BPST. General students' performance, as well as their performance in individual science process skills, was analyzed through descriptive statistics. Analysis of variance (ANOVA) and independent samples t-test, on the other hand, were used to statistically determine whether there was the difference in the performance of students in the specific process skills based on their sex and grade levels .

5.3 RESULTS AND DISCUSSION: THE KNOWLEDGE LEVEL OF SCIENCE PROCESS SKILLS OF SECONDARY SCHOOL STUDENTS IN TANZANIA

5.3.1 Introduction

This section presents and discusses the key findings of the study to examine the level of science process skills of Tanzania students using the validated test instrument (BPST). Advanced level Biology students in Morogoro municipality schools were taken as a case study. The section is subdivided three major parts depending on the questions being addressed. The following research questions guided this research phase. Section 5.3.2 is the general performance of advanced level Biology students in science process skills test in the selected schools in Morogoro municipality. Section 5.3.3 is the comparison of the performance of advanced level Biology students in Morogoro municipality based on their demographic characteristics of sex and grade level. The performance of Biology students each of the five process skills (formulating hypotheses, defining operationally, controlling variables, design experiments and interpreting data) are discussed in section 5.3.4. The last part of the section presents summary of findings and some conclusions.

5.3.2 General performance of Morogoro students in the Biology process skills test (BPST)

The first objective of this stage was to assess the general knowledge level of integrated science process skills of Morogoro Biology students by using a science process skills test developed in the first phase. The test was administered to a group of 353 advanced level Biology students from all four Biology based schools present in the municipality of Morogoro. The study involved 246 (69.7%) female students and 107(30.3%) male students of which 142(40.2%) were form six students and 211 (59.8%) were form five students. Descriptive statistics was performed to examine means, standard deviations, percentages, and frequency distributions of scores. Descriptive statistics indicates that the mean score of students was 17.2 (49.1%) with s.d of 7.3. The highest score was 28 (80%) and the lowest 09 (25.7%) out of 35 possible. 66 (18.6%) students out of 353 scored 18 (51.4%) out of 35 and this was the mode score, followed by 15 (42.8) which was scored by 54 (15.2%) of all students who participated in the study. More statistics descriptive to the general performance of science process skills is given in Table 5.4. According to the table, (Table 5.4) majority of Morogoro Biology students 116 (32.8%) out of 353 scored average on the scale grade ie between 14-17 out of 35 maximum possible and were classified as having an average performance . Some 99 (28%) students out of 353 in the sample had satisfactory knowledge level of process skills (18 – 20) as shown in Table 5.4 below.

Table 5.4 Descriptive statistics of student scores in the BPST instrument (n=353)

Range of scores by %	Correspo-nding %	Grade	No. of students in the range	% of students in the range	Description of the level of process skills
0- 6	0-19	F	0	0.0	Very unsatisfactory
7 – 10	20-29	E	6	1.7	Unsatisfactory
11- 13	30-39	D	43	12.2	Below average
14- 17	40 - 49	C	117	33.1	Average
18 – 20	50- 59	B	99	28.0	Satisfactory
21- 26	60-74	B+	86	24.3	Very Satisfactory
27 – 35	75-100	A	2	0.6	Excellent

Source: Field data (2014).

Although the table shows that none (00%) of the students had F grade, implying that all Biology students scored more than 06 items out of 35. However only 02 (0.6%) students out of 353 scored A grade. These excellent graded students both scored 27 out of 35 possible which is equivalent to 77%. Some 06 (1.7%) students scored between 7– 10 items and they were graded as unsatisfactory while 43 (12.2%) scored below average 30-39%. On the other hand, 86 (24.3%) students scored between 60-74% (21-26) and from the secondary education grading system of Tanzania, they were graded as having a very satisfactory level of science process skills. Skewness of scores which is the extent to which a distribution of values deviates from symmetry around the mean was also calculated and a value of 0.046 was obtained. A value of a positive 0.046 skewness means that there were a relatively greater number of smaller values than mean (Dover, 1979). It also indicates that most of the students taking the test obtained low scores. On the other hand, the overall mean score was 17.2 (49.1%) which means that on average, the advanced level Biology students in Morogoro scored between 17 to 18 items correctly out of 35 total questions.

According to the grading system of Tanzania adopted in this study, 49.1% represents a “C” class which means average knowledge level. This means that on overall, Morogoro Biology students have barely average knowledge level of integrated science process skills.

Their overall level of performance in this study cannot be regarded as "good" considering the emphasis placed on the acquisition of science process skills in their new science curricula. Some of the possible reasons for the students' weak performance might be that many might not be familiar with the types of tasks investigated and assessment used in this study. Germann et al. (1996) asserted that students' good performance on science process skills was dependent on their experience with and domain-specific practice activities on the skills in prior tasks, while Ruiz-Primo & Shavelson (1996) reported that student scores depended on the particular tasks investigated and on the particular method used to assess their performance. In the similar vein, Millar & Driver (1987) found that students' ability to use process skills depend on the extent of their knowledge of the contexts they are asked to work on. This is also explained by the finding (Rowe & Foulds, 1996; Tobin & Capie, 1982) that performance of tasks requiring these process skills is strongly content-dependent. There is a problem of how to integrate content and process of science in Tanzania. Science process skills exercised in relation to some science content and have a crucial role in the development of learning with understanding (Harlen, 1999). Tanzania science teachers need to capitalize on opportunities in the activities done in the science classroom to emphasize science process skills. Students conducting these activities are expected to develop such skills as stating hypotheses, operationally defining variables, designing investigations, and interpreting data in addition to mastering the content of the courses.

This might be due to the fact that the conventional methods, that predominates in Tanzania science classrooms of all levels (Osaki et al. 2004), does not facilitate the development of generalizing skills and other science process skills in the subjects. Osaki (2007) attributed this to poor science teacher preparation in teacher training institutions. According to the author, teacher education curriculum has failed to promote reflective practices and constructivist approaches to prospective science teachers. As a result, these institutions are increasingly producing teachers who are weak in practical skills especially laboratory experiences (Osaki, 2007).

The finding from this study however, differ from the findings by Beaumont-Walters & Soyibo (2001) who conducted a study to determine Jamaican high school students' level of performance on five integrated science process skills and if there were statistically significant differences in their performance linked to their gender, grade level, school location, school type, student type and socio-economic background (SEB). Data collected with the authors' constructed integrated science process skills test the results indicated that the subjects' mean score was low and unsatisfactory; their performance in decreasing order was: interpreting data, recording data, generalizing, formulating hypotheses and identifying variables; there were statistically significant differences in their performance based on their grade level, school type, student type, and SEB in favor of the 10th graders, traditional high school students, ROSE students and students from a high SEB. However there was a positive, statistically significant and fairly strong relationship between their performance and school type, but weak relationships among their student type, grade level and SEB and performance (Beaumont-Walters & Soyibo, 2001).

5.3.3 Performance of Morogoro Biology students based on their demographic variables of sex and grade level

The second purpose of this stage of the study was to determine if there were any significant differences in the students' performance on the BPST linked to their demographic variables of gender and grade level. An independent samples t-test was employed to estimate the accounted variances of the students’ science process skills based on grade level and sex. Their means and standard deviations were computed and subjected to computer SPSS independent samples t-test as shown and discussed in the following sub-sections below.

5.3.3.1 Performance of Morogoro Biology students based on their sex

Means (averages) of students test scores based on sex were computed and subjected to the independent samples t-test. The study involved 107 (30.3%) male and 246(69.7%) female advanced level Biology students. As it has been summarized in table 5.5 (a) below the mean of scores for male students was 15.76 (45.1%) and the standard deviation of 7.10 while the mean for female students was 17.75 (50.7%) and the standard deviation of 7.32. Descriptively, it means that female students outperformed their male counterparts in the Biology process skills test (BPST). The finding that the females outperformed the males in the science test was surprising. This is because it conflicts with many previous studies' findings for example studies by Forrest, 1992; Gladys, 2001; Greenfield, 1996; Johnson, 1987; and that by Klein et al. 1997.

The statistically significant difference was also found between student scores based on their sex when null hypothesis was subjected to independent samples t-test. The null hypothesis stated that "there is no statistically significant difference in the performance of Morogoro Biology students based on their sex in the Biology process skills test". An analysis of independent samples t-test based on sex at alpha (α) =0.05 produced a p of 0.019 and a t value of 2.363, hence to reject the stated null hypothesis. As indicated in tables 5.5 (a and b) below, a t-test revealed a statistically significant difference between the mean score of male students (M = 17.75, s = 7.32) and that of female students (M = 15.76, s = 7.10), t (351) = 2.363, p = 0.019, α = 0.05. Hence, the null hypothesis that there is no a statistically significant difference in the performance of Biology students based on their sex in the science process skills test was rejected at 0.05 alpha. This implies that Morogoro female students statistically outperformed their male counterparts in science process skills test. Tables 5.5 (a and b) below summarize the independent samples t-test of scores based on students' sex.

Table 5.5 (a): *Group statistics t-test for test scores based on sex (n=353)*

	Sex of the respondent	*N*	*Mean*	*Std. Deviation*
Respondent level of integrated process skills	*Male*	*107*	15.76	7.10
	Female	*246*	17.75	7.32

Source: Field data (2014).

Table 5.5 (b): *Independent samples t-test for test scores based on sex (n=353)*

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	T	Df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Students level of process skills	Equal variances assumed	0.169	0.681	2.363	351	0.019	1.98568	0.84017	0.33328	3.63808
	Equal variances not assumed			2.391	207.267	0.018	1.98568	0.83048	0.34841	3.62295

Source: *Field data (2014).*

The findings that Morogoro female students statistically outperformed their male counterparts in science process skills test differs with the result obtained by Dyer & Myers (2006) who investigated on how male and female students integrated laboratory experiences on student content knowledge and science process skills. They found that the sex of the students did not contribute significantly to the variance in content knowledge achievement. In their findings on the effect of gender on students problem-solving skills, Shaibu & Mari (1997) discovered that female students were significantly better than their male counterparts in the ability to solve quantitative problems. Trigwell (1990), Bazler & Simonis (1991), and Baker & Leary (1995) have also reported that female students have more difficulties than male students in chemical problem-solving topics. On the other hand, Morogoro result somehow differs from those by (Riggs, 1991), who conducted a study to examine how gender affects elementary-school students attitudes and ability to learn sciences. The results using 331 students revealed a significantly higher score for males than females. The finding that there was a significant gender difference in their performance is also inconsistent with the finding of some other researchers such as Greenfield (1996) and Walters and Soyibo (2001). The researcher of the current study could not be able to explain the likely reason for this finding based on this study's data.

5.3.3.2 Performance of Morogoro Biology students based on their grade

Another purpose of this study was to determine if there is a statically significant difference in the students' performance on the sciences processes skills test linked to their grade level. Independent samples t-test involving the means of Form V and Form VI scores was applied for analysis. The study involved 211(59.8%) Form V and 142 (40.2%) Form VI Biology students studying in the municipality of Morogoro. As it has been summarized in table 5.6 (a) below the mean of scores for Form V students was 17.00 and the standard deviation of 7.98 while the mean for Form VI students was 17.37 and the standard deviation of 6.81. Table 5.6 (a) shows that there is a statically difference in the students' performance on sciences processes skills test linked to their grade level ($p<0.05$). As seen in table 5.6 (a) below, the mean of scores of the Form VI students was higher than their counterpart Form V before the scores being subjected to the independent samples t-test.

Table 5.6 (a) *Group statistics t-test for test scores based on grade level (n=353)*

	Grade level of the students	*N*	*Mean*	*Std. Deviation*
Grade level of the students	Form V	211	17.00	7.98
	Form VI	142	17.37	6.81

Source: Field data (2014).

However, no statistically significant difference was found between student scores based on their grade level when null hypothesis was subjected to independent samples t-test. The null hypothesis stated that "there is no statistically significant difference in the knowledge level of Morogoro Biology students based on their grade level". An analysis of independent samples t-test based on grade level at alpha (α) =0.05 produced a p of 0.638 and a t value of 0.470, hence, it failed to reject the null hypothesis as indicated in Tables 5.6 (b) below. A t-test failed to reveal a statistically significant difference between the mean score of Form V students (M = 17.00, s.d = 7.98) and that of Form VI students (M = 17.37, s.d = 6.81), t (351) = 0.470, p = 0.638, α = 0.05. Hence, the null hypothesis, that there is no statistically significant difference in the scores of Morogoro Biology students in the BPST based on grade levels was accepted at 0.05 alpha levels. Tables 5.6 (b) below summarize the independent samples t-test of scores based on students grade level.

Table 5.6(b) Independent samples t-test for test scores based on grade level (n=353)

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	T	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Students level of process skills	Equal variances assumed	2.084	0.150	0.470	351	0.638	0.37324	0.79347	-1.18732	1.93380
	Equal variances not assumed			0.456	269.743	0.649	0.37324	0.81821	-1.23765	1.98413

Source: *Field data (2014).*

The table shows that there is a no statically difference in the students' performance on the Biology processes skills test linked to their grade level (p<0.05). That means although descriptively the mean of scores of Form VI students seems higher than their counterpart Form V, but statistically, there is no difference. This finding receives some indirect support from Gallagher (1994) regarding middle school students' performance on science process skills and the finding of many previous studies on the link between students' economic background and science performance (Blosser, 1994).

On the overall, this study was framed conceptually from the reported findings that certain students variables notably years of learning experience and sometimes sex are important determinants their competence in science related disciplines. Anecdotal evidence would suggest male students at higher grade levels (many years of the learning experience) have higher knowledge level of integrated science process skills than students who miss these qualities. However, that was not the case with Morogoro Biology students in this study. Statistical tests of the null hypothesis in this study found a statistically significant difference in terms of the knowledge level of science process skills of students in Morogoro based on their sex in favor of female students. Female students statistically outperformed their male counterparts in the scientific skills test.

Statistical tests of a null hypothesis also found that regardless of their grade level (Form V vs Form VI) there was no statistically significant difference in terms of their knowledge of science process skills. This means that the grade level of the student had no significant influence on the knowledge level of science process skills. It can be concluded that regardless of years of learning experiences (grade level) and sex Morogoro students possesses unsatisfactory knowledge level of science process skills as seen in section 5.3.2. This finding implies that the teaching of content must take precedent over the training of students in science process skills. Successfully integrating these skills into classroom lessons and field investigations will make the learning experience richer and more meaningful.

5.3.4 Performance of Biology students by specific science process skill

5.3.4.1 General overview of students' performance by specific science process skill

Another objective of this stage was to examine the performance of advanced level Biology students of Morogoro municipality based on the five integrated science process skills namely i. formulation of hypotheses ii. identifying and controlling variables, iii. design experiments iv. analyzing and interpreting data and v. defining variable operationally. Therefore for additional analyses, the entire Biology process skills test down into its five subscales of process skill objectives. The mean scores end standard deviations on the BPST total and each subscale and overall students were calculated and summarized in Table 5.7 below. For overall students, correct response percentages were highest for the process skills of identifying and controlling variables with the mean score of 4.05(57.7%) and were lowest for the process skills of analysis and data interpretation with the mean score of 2.34(33.4%). The mean of a raw score of the subtest was low, indicating that the students found the subtest more difficult.

Table 5.7 below is a summary of descriptive statists and the difficulty levels of each process skill objective as a subscale of the BPST.

Table 5.7 Descriptive statistic of the test and its component skills (n=353)

Specific science process skills	Total items	Minimum Score	Maximum Score	Mean Score	SD	Percent Correct
Identifying and controlling variables	7	1	6	4.05	0.88	57.8
Identifying and stating hypotheses	7	1	6	3.49	1.43	49.8
Operationally defining	7	1	6	3.71	0.96	53
Analysis and interpreting data	7	0	5	2.34	0.75	33.4
Designing experiments	7	1	6	3.27	0.96	46.7

Source: Research survey (2014)

As it is shown in table 5.7 above, Morogoro Biology students had a mean or average of correct responses of 3.71 in questions measuring their skills in defining variables operationally. Student scores in questions measuring this skill ranged from 01 as the lowest score to 06 out of 07 questions as the highest score. Students, on the other hand, had a mean of 3.49 (an average performance) in questions measuring their skills in hypothesis formulation. In these questions, student scores ranged from zero (01) out of six to 06 out of seven (07) items. Lastly, Morogoro students had a mean of 3.27 (average of performance) on questions measuring their skills in designing experiment. In these questions, teachers had scores ranging from zero (01) to five (06) out of seven questions.

The table shows that subjects did perform relatively better on the skill of identifying and controlling variables probably because most of the items requiring this skill gave prescriptive directions on what the subjects should measure and how to record (first level of the developmental progression of the skill). But a close look at the subjects' test scripts revealed that only of a few of them were able to interpret tables and graphs and record data in more complex table form on their own and that they were better able to complete and construct tables than graphs. The construction of graphs demands the ability to recognize relations between relations or formal operations in Piagetian terms which many students are incapable of (Shayer & Adey, 1981). The subjects performed fairly well on interpreting data that demanded extracting information from graphs and tables, but they were less successful (barely "average")on the skill of generalizing which entailed making conclusions,interpolating/extrapolating between/beyond data points and identifying supporting evidence.

Performance of Morogoro Biology students by specific integrated science process skills somehow resembles findings reported by both Hamilton & Swortzel (2007) and Dyer et al. (2004) where students scored higher on questions measuring their skills in identifying variables and stating hypotheses and also scored poorly on measuring their ability in graphing and data interpretation. As it is shown in Table 5.7 above, Morogoro Biology students performed poorly on questions dealing with analysis and interpretation of data with the performance mean of only 2.34 or 33.4% correct responses. The maximum score out of seven items was only 5 and the largest standard deviation of 2.024. Morogoro students result, however, correspond with those by Hackling & Garnett (1991) who conducted a research on students ability in carrying out experiments and found that students at all levels showed a poorly developed skill of problem analysis, planning, and carrying out controlled experiments. Another similar finding is that by Foulds & Rowe (1996) who found that students were capable of identifying all variables influencing an experiment, scoring about 50% on the test items and they could also produce testable hypotheses, with scores of about 40%. However, they were unable to design a controlled experiment and analyze experiment results, gaining an average mark of only 18%. The students 'poor performance on the skills of analyzing and interpreting data might be due to the likelihood that they had not been taught these skills and that their levels of cognitive development were inadequate to enable them to handle the skills. It is in the view of this study the teacher-centered mode of teaching science in the sampled schools, which did not allow the Biology students to practice and internalize the skills over a fairly long period, was likely to be one of the main reasons for the students' poor performance on the skills.

5.3.4.2 Detailed description on the performance students by specific science process skills under the study

The study also intended to provide full description of the performance of Biology students in each of the five integrated science process skills namely i. formulation of hypotheses ii. identifying and controlling variables, iii. design experiments iv. analyzing and interpreting data and v. defining variable operationally. The mean of students' scores and standard deviations on each subscale were calculated. The following section discusses the performance of Morogoro Biology students in each scientific skill focused by this study.

5.3.4.2.1 Performance of students in the skill of control of variables skill

One aspect of the inquiry practice that directly related to student ability to carry out inquiry-oriented investigations is the ability to handle and control experimental variables. Control of variables as a fundamental science process skill has been widely regarded as an important ability in scientific investigations and as an integral component of most curricular around the world (Turaib, 2015). For overall students, correct response percentages were highest for the process skills of identifying and controlling variables with the mean score of 4.05(57.7%) out of seven items measuring this skill (see table 5.7 section 5.3.4.1 above). Although the percentage of students who showed understanding of the concept of control of variables represents less than two-third of the sample, it is still fair better than when compared to the performance in other subscales. However, during marking their tests, it was observed that students were not able to tell whether a particular variable influenced or determined the results of the experiment. This means much work is needed to improve students' ability to handle and control experimental variables into Tanzania science learners.

Across many studies, it is evident that most students and even some adults do not have a generalized understanding of controlling variables because of their ability to identify, select, or design controlled experiments depends on the task content or situational factors (Koslowski, 1996; Linn et al. 1983; Zimmerman, 2000). This skill provides students with the scope and understanding needed to carry out controlled and reliable experiments that might eventually lead to trusted outcomes and valid inferences (Chen & Klahr, 1999). The findings from this study are in congruence with the finding obtained in the study by Turaib (2015) in his study to assess students' understanding of the control of variables across three grade levels and gender in the United Arab Emirates (UAE). His findings revealed that students across grade levels exhibited alternative conceptions of key ideas related to control of variables. Similar findings have also been seen by Boudreaux et al. (2008) who found that although most of the students participating in their study were able to realize the importance of having controlled conditions for experimentation, many students had difficulties in providing a valid justification for why controlled conditions were important. Research studies in this area call for critical investigations to suggest and develop methods and approaches needed to help students develop sound and coherent understanding of this crucial and essential skill (Zimmerman, 2000). The findings with Morogoro students highlights the need for teachers to pay attention to the development of argumentation and analytical skills needed to argue for which variables need to be manipulated and which ones need to be kept constant. A suggested by Turaib (2015) students need to focus on simple steps of recognizing variables of experiments and categorize them into categories so that decisions about their manipulations can be made.

5.3.4.2.2 Performance of Morogoro students in the skill of data analysis and interpretation

Data analysis entails the ability of students to assign meaning to the collected information and determining the conclusions, significance, and implications of the experimental findings (Zimmerman, 2007). Analysis of BPST scores indicated that students' scores were lowest for the items measuring their ability in data analysis with the mean score of only 2.34(33.4%) out of seven (07) items (see table 5.7. section 5.3.4.1 above). Compared other subscales, data analysis had the smallest standard deviation of 0.75. This implies that Morogoro students were so hemogeneous in terms of their ability in data analysis questions and that many students had scored the same scores as their mean score. These findings that Morogoro students had poor scores in data analysis resembles the findings reported by both Hamilton & Swortzel (2007) and Dyer et al. (2004) where students scored higher on questions measuring their skills of controlling variables but scored poorly on items measuring their ability in graphing and data interpretation. These finding on Morogoro students also correspond with those by Hackling & Garnett (1991) who conducted a research on students ability in carrying out experiments and found that students at all levels showed a poorly developed skill of problem analysis, planning, and carrying out controlled experiments. Another similar finding is that by Foulds & Rowe (1996) who found that students were capable of identifying all variables influencing an experiment, scoring about 50% on the test items and they could also produce testable hypotheses, with scores of about 40%. However, they were unable to design a controlled experiment and analyze experiment results, gaining an average mark of only 18%. The complexities surrounding understanding of the concept of data analysis extend to science teachers. In an early study, Shadmi (1981) studied science teachers' understanding of the control of variables and found that most teachers had difficulty interpreting the results in the context of experimental settings.

The poor students´ performance on the skills of analyzing and interpreting data might be due to the likelihood that they had not been taught well enable them to handle this skill. It is in the view of this study that, teacher-centered model of teaching science in the sampled schools in Morogoro, did not allow the students to practice and internalize the skills over a fairly long period. This is likely to be one of the main reasons for the students' poor performance on the skills. This means that current teaching-learning processes should not only focus on conceptual understanding of science, but it must also move in directions similar to those identified in science education research as 'doing science' and 'knowing about science' (Zimmerman, 2007). In order to achieve this goal, teaching and learning processes must focus on equipping students with the intellectual and the manipulative skills that are needed to construct and reconstruct scientific knowledge rather than focusing on conceptual learning only.

5.3.4.2.3 Performance of Morogoro students in the skill of formulating hypotheses

A hypothesis is an educated prediction that can be tested. Formulating hypotheses is a scientific way in which the investigator forms a research hypothesis that states an expectation to be tested. Then the investigator derives a statement that is the opposite of the research hypothesis. This statement is called the null hypothesis (H_0) (Ghanem, 2003). This study also intended to determine the knowledge level of students in formulating and stating testable hypotheses. The findings from BPST indicated that Morogoro students scored below average on the items measuring their ability in formulating a hypothesis. As seen in table 5.7 section 5.3.4.1 above, the mean of seven items measuring their ability in this skill was 3.49(49.8%) and the standard deviation was 1.43. Student scores ranged from one (01) to six to 06 out of seven (07) items.

These findings that Morogoro students have below average performance in items measuring their hypothesis formulation skills were not surprising. Many researchers who have studied hypotheses formulation within science education have concluded that students have weak abilities in formulating and testing hypotheses. According to Ghanem (2003) students incur three main problems when dealing with scientific hypotheses. These problems include failure to formulate valuable examined hypotheses; failure to distinguish between scientific facts, theories, and hypotheses, and difficulty in verifying hypotheses. For example, in their study on young children differentiation of hypothetical beliefs from evidence, Sodian et al. (1991) found that students tend to produce or repeat the effect rather than to discover its causes and they have trouble on identifying likely causes. Furthermore, students were unable to quickly grasp the meaning of the investigated subject, method, and the image of solving the problem (Sodian et al. 1991). The findings of the current study however, highlight the fact that better preparation of students for the future may require new teaching approaches that respond to and focus on not only learning scientific content but also on acquiring transferable abilities such as the ability to design and conduct valid and controlled experiments that yield valid and reliable findings. As the observation made by Filson (2001) that students have difficulty with hypothesis because their books and lessons mention hypothesis, but almost never really explain or model them and frequently hypotheses are confused with theories.

5.3.4.2.4 Performance of Morogoro students in the skill of designing scientific experiments

Developing the ability to design an experiment is critical to the understanding of the scientific process and in promoting critical thinking skills (Coil et al. 2010). Ths study also measured students´ knowledge level of designing experiments scientifically using BPST. Analysis of students score in this subscale indicated that Biology students had also a below average ability in designing experiments. The mean score of students in this subscale was 3.27(46.7%) while the standard deviation was 0.96 (see in table 5.7 section 5.3.4.1 above). These findings that Morogoro Biology students have below average performance in items measuring ability in designing experiments were also not surprising. A number of science education researchers (Coil et al. 2010; Chen & Klahr, 1999; Adey & Shayer, 1990; & Ghanem, 2003) attribute poor students' ability in correctly designing experiments to misconceptions and inaccuracies regarding randomization, sample size, and inability to identify and control variables and poor stated hypotheses. According to Adey & Shayer (1990) students weak in designing experiments because they are rarely given an opportunity to think deeply about experimental design or asked to develop experimental protocols on their own.

Scores from BPST showed that most of Morogoro students know that an experiment should contain a control, but many find it difficult to define exactly what a control is. Similar observation was made by Klymkowsky et al. (2011) in their study which intended to reveal student thinking about experimental design and the roles of control experiments. In this study Klymkowsky et al. (2011) surprisingly found that a high percentage of students had difficulty identifying control experiments even after completing three university-level laboratory courses. To address this problem Klymkowsky et al. (2011) designed and ran a revised cell biology lab course in which students participated in a weekly experimental control exercise. Not unexpectedly, the results indicate that the revised course led to greater improvements in students' ability to identify and explain the purpose of control experiments. So it can be concluded that using a simple experimental measure, students can become engaged in the process of scientific inquiry, and in turn, begin to think deeply about experimental design. This skill can be developed if students are allowed to work like scientists.

5.3.4.2.5 Performance of Morogoro students in the skill of defining operationally

Defining operationally means developing statements that present concrete descriptions of an event by telling someone what to do or what to observe (Chiappetta & Koballa, 2002). It is a specific definition of a concept in a research study. Another specific aim of this study was to measure the knowledge level of Morogoro Biology students in defining variables operationally. It has to be noted that, once researchers develop hypotheses, the next step involves forming operational definitions of the concepts to be investigated in the research (Klymkowsky et al. 2011). So it is one of the very vital integrated science process skill to be acquired by students. Analysis of students score in this subscale indicated that Biology students had above average ability in defining terms operationally. As seen in table 5.7 section 5.3.4.1 above, the mean score of students in this subscale out was 3.71(53%) out of seven items which measure this skills. The standard deviation was 0.96.

Few studies exist which explains how a teacher can help students define experimental variables operationally. Pratt & Hackett (1998) suggest that, by learning science through inquiry, a science teacher can facilitate the development of defining operationally skill and acquisition of science process skills in general. Teachers are taught inquiry teaching strategies by engaging in inquiry science activities and extending their understanding of the science concepts that they teach (Hyman & Shephard, 1980). According to Harlen (2000), teachers can facilitate the development of defining operationally skill and other science process skills in general by; (i) providing a variety of materials and resources to facilitate students' investigations, (ii) posing thoughtful and open-ended, (iii) encouraging dialogue among students and with the teacher, and (iv) keeping students' natural curiosity alive during teaching. Nevertheless it was enough for this study to indicated that Biology students had above average ability in defining terms operationally compared to other subscales.

5.3.5 Summary of findings on Students performance in the BPST

Science process skills, as in the Tanzania´s competence based curriculum of 2005, have been identified in the science education literature as an effective inquiry method of teaching science. This study aimed at assessing the knowledge level of advanced level Biology students in the municipality of Morogoro of science process skills. Based on the Biology process skills test (BPST) scores, it was found that Biology students in Morogoro municipality had barely average knowledge level of integrated science process skills. The mean of test scores was 17.2 items out of 35 items in the test corresponding to 49.1%. However, Morogoro students performed relative better on items measuring their ability in identifying and controlling variables with score mean of 4.05 out of 07 items and they performed extremely poor on items which measured their skills in analyzing and interpreting data with the mean of 2.34 out of 07 items.

Due to the influence of social forces, culture and gender roles in the Tanzania, anecdotal evidence would suggest male students to have higher levels of achievement in science-related disciplines than females. However, the findings from Morogoro biology students in this study did not support that assertion. Based on the science process skills test scores of the 246 females and 107 males in the study, female had a relatively higher mean of 17.75 than their counterpart male who had a score mean of 15.76. Statistical significant differences were found to exist between male and female students in terms of their performance in science process skills through SPSS t-test. Although experience and maturity might be strong determinants for one's academic performance, an independent t-test of students score means based on their years of schooling in this study failed to find a statistically significant difference of Form V and Form VI students. Form VI has one extra year in the education system than their counterpart Form V. The mean of 142 Form VI who participated in this study was 17.37 while the mean of 211 Form V was 17.00.

CHAPTER SIX

THIRD STAGE: AN INTERVENTION STUDY TO INVESTIGATE THE EFFECTIVENESS OF AN INQUIRY-BASED APPROACH *VS* CONVENTIONAL METHOD IN THE DEVELOPMENT OF STUDENTS' SCIENCE PROCESS SKILLS, CONCEPTUAL UNDERSTANDING, AND MOTIVATION

6.1 Introduction

This third stage of the study aimed at investigating the effectiveness of an inquiry-based approach on students' process skills development, conceptual understanding and motivation by compairing it with the traditional mode of teaching. Genetics was taken as a case study and the motivational constructs focused in the study included i. intrinsic motivation, ii. grade motivation, iii. career motivation, iv. self-efficacy, v. self-determination, and vi. self-concept. Inquiry-based approaches to science have been heavily emphasized by the newly adopted competence based curriculum of 2005 in Tanzania. The curriculum encourages science teachers to move from traditional and conventional methods and use participatory inquiry-based approaches as much as possible. However, no study has been conducted to assess the effectiveness of the suggested approach especially in students' science process skills development, conceptual understanding of contents, and motivation to science. Inquiry-based approaches follow the moderate constructivism learning theory which emphasizes students to construct their own understanding and knowledge of the world, through experiencing things and reflecting on those experiences (Eiskenkraft, 2003).

In order to compare the effectiveness of the inquiry-based approach against the conventional method, six months were spent during the summer of 2015 in teaching themes within genetics at three selected schools in the vicinity of Morogoro municipality. The study employed a quasi-experimental research design with pre and posttests (before and after the teaching intervention). Eight (08) weeks genetics teaching course were designed on the basis of both (i) inquiry based learning principles and (ii) conventional lecture style.

i. For inquiry genetics course, free online inquiry Genetics lessons, which are in line with the Tanzania syllabus for advanced level students were collected from different websites (see Table 6.2 and in section 6.2.4.1).

ii. Books prescribed by the Tanzania Biology syllabus were used in the preparation of 8 weeks conventional genetics lessons.

The intervention was implemented into Form V and Form VI Biology classes of the three selected secondary schools in Morogoro. Secondary schools that were involved in this study included Kilakala (145 students), Alfagerms (87 students) and Bigwa sisters (31 students). Form six classes were taught using conventional method while form five classes in these schools had enough time and were taught using inquiry approach. Both classes had never been exposed to advanced level genetics. Section 6.2 of this chapter explains the methodological aspects that were employed by the researcher in this stage while section 6.3 presents and discusses key findings of this quasi-experimental study on the effectiveness of inquiry approach to science. On the otherhand, the discussion on the correlation between students achievement in science process skills with their conceptual understanding and motivation is presented in section 6.4.

6.2 RESEARCH METHODOLOGY OF THE THIRD STAGE

6.2.1 Research design

Both correlational and quasi-experimental designs were employed at this stage. Correlational research involves the search for relationships between variables through the use of various measures of statistical associations such as Chi-square, Student's t and F tests (Borg & Gall, 1989). Correlational design technique was employed here because the study partly aimed at exploring the relationship between students' competence in science process skills with their conceptual understanding of contents and motivation towards science process skills. The interest was only whether or not correlations of these variables exist and not causality. However, a large part of this intervention study was quasi-experimental involving experimental and control groups. This is because secondary school classes exist as intact groups and school authorities do not normally allow classes to be dismantled and reconstituted for research purposes (Shadish, Cook and Campbell, 2002; Njoroge et al. 2014). Hence there was a non-random assignment of students to the groups. Quasi-experimental researches are widely used in the evaluation of teaching interventions because it is not practical to justify assigning students to experimental and control groups by random assignment (Randolph, 2008; Njoroge et al. 2014). Quasi-experimental research offers the benefit of comparison between groups because of the naturally occurring treatment groups (Cohen et al. 2007). In this study, the experimental groups were exposed to the treatment (inquiry-based approach) and the control groups received no treatment (they were taught using traditional methods only). The pre and posttests included Biology process skill test (BPST), Genetics conceptual understanding test, self-concept Likert scale by Damerau (2012) and science motivation measuring scale (SMQ-II) by Glynn et al. (2011). The performances of the two groups were then compared to determine whether there are any treatment effects as a result of different teaching styles on the same contents.

6.2.2 Data Collection Methods

In order to collect as much information as possible for this study, several data collection tools were either constructed or adopted depending on the specific objective under question. The following data collection instruments were used,

i. Test of science process skills specific to Biology (BPST)
ii. Science Motivation Questionnaire II by Glynn et al. (2011)
iii. FSWEx Self-concept scale by Damerau (2012)
iv. Genetics test -Cognitive test for measuring conceptual understanding of contents of genetics

6.2.2.1 Science Motivation Questionnaire by Glynn et al. (2011)

The Science Motivation Questionnaire-II by Glynn et al. (2011) which is a five-point scale Likert-type questionnaire was employed as one of the data collection tools at this stage for collecting students' affective characters before and after intervention (see appendix V). The Science Motivation Questionnaire-II was developed to enhance the construct validity of the Science Motivation Questionnaire (Glynn et al. 2011). It examines how motivated students are to learn science, and why those who are not motivated feel that way (Glynn et al. 2011). According to Glynn et al. (2011), Science Motivation Questionnaire-II (SMQ-II) is a 25 item scale and assesses five components of students' motivation to learn science in college or high school courses. The five components of motivation assessed include i. intrinsically motivated science learning, ii. grade motivated science learning, iii. self-determination for learning science, iv. confidence (self-efficacy) in learning science, and v. career motivation for learning science (Glynn et al. 2011). It was the intention of this research to study the influence of inquiry-based instruction in these motivational components.

According to Glynn et al. (2011) the SMQ-II has an internal consistency reliability coefficient of $\alpha = .92$. Glynn et al. (2011).They also found content validity evidence through expert review of the items as well as construct validity evidence through exploratory and confirmatory factor analysis. Further evidence for the scale's construct validity was shown in that science majors scored higher on all of the scales, with effect sizes ranging from $d = 0.33$ to $d = 1.17$. Finally, they confirmed criterion-related validity evidence through correlating the results of the scale with performance indicators, such as the students' science GPA. Students' total scores significantly correlated with reported high school preparation in science, college science GPA (r¼0.56, p<0.01), and the relevance of science to their careers (r¼0.56, p<0.01) (Glynn et al. 2011). The items are strongly worded, unambiguous declarative statements in the form of short, simple sentences without jargon and easy to read. The maximum total score of the scale is 125 and the minimum is 25. Students who score from 25 to 49 are ‘‘never to rarely'' motivated, 50–79 are ‘‘rarely to sometimes'' motivated, 80–109 are ‘‘sometimes to often'' motivated, and 110–125 are ‘‘often to always'' motivated. The scale is usually self-administered, as part of a more comprehensive questionnaire. A student requires 10-18 minutes to complete it. According to Glynn et al. (2011), SMQ-II might become Biology Motivation Questionnaire, Chemistry Motivation Questionnaire (CMQ), or Physics Motivation Questionnaire (PMQ) in which the words Biology, Chemistry, and Physics are respectively substituted for the word science. The motivational components and their associated items include intrinsically motivated science learning (items 1, 3, 12, 17 and 19), grade motivated science learning (items 2, 4, 8 20 and 24), self-determination (responsibility) for learning science (items5, 6, 11, 16 and 22), self-efficacy (confidence) in learning science (items 9, 14, 15, 18 and 21), and career motivation (items7, 10, 13, 23 and 25) (Glynn et al. 2011).

Students responded to each of the 25 randomly ordered items on a five-point Likert-type scale of temporal frequency ranging from 0 (never) to 4 (always).

6.2.2.2 Biology process skills test (BPST)

In assessing the knowledge level of integrated process skills of advanced Biology students in Morogoro, the Biology process skills test (BPST) developed and validated in the first stage of this study was used (see Appendix I). The test measures five (05) individual integrated scientific skills (identifying variables, stating hypotheses, operationally defining, designing investigations and analyzing and interpreting data) to advanced secondary school learners. The reliability of the instrument was established by the researcher in the year 2014 using 610 learners to be 0.80 (Cronbach's alpha). Concurrent validity of BPST was established by comparing students score in the process skills test (TIPS II) by Burns et al. (1985) and found to be 0.51. Using experts' opinion scale, the content validity of BPST was found to be 0.88. The test has reliability coefficient well above the lower limit of the acceptable range of values for reliability, and it is within the range of reliability coefficients obtained from similar studies, such as those by Dillashaw and Okey (1980) who obtained a reliability of 0.89 and Burns et al. (1985) who also obtained a reliability of 0.84. Biology process skills test (BPST) has a readability index of 72.2. This high readability value implies an easy to read text to students who English is not their first language like Tanzania students. The researcher adopted this test because it has been developed in the context of Tanzania using the Tanzania competence based curriculum.

6.2.2.3 FSWEx self-concept scale by Damerau (2012)

In assessing the level of students' self-concept towards science process skills before and after teaching intervention, the FSWEx self-concept scale by Damerau (2012) was used (see appendix VI). This questionnaire was designed to enable researchers and science teachers to gain a better understanding of the self-concept of students in doing science and to examine in which ways it affects the interest of doing science. FSWEx Self-concept questionnaire consists of 18 items which are further subdivided into three subscales, i. planning experiments (06 items), ii. practical experimentation, (06 items) and iii. analyzing data (06 items).

The scale is based on the model of experimental skills (Schreiber et al. 2009) and uses a 5-point Likert scale ranging from 0 (strongly disagree), 1 (disagree), 2 (neutral), 3 (agree) to 4 (strongly agree). The internal consistency reliability of the instrument (FSWEx Self-concept scale) was established by Damerau (2012) using 177 grade 11 to 13 science learners to be $\alpha = 0.77$ (Cronbach's alpha). The three subscales of FSWEx i. planning experiments, ii. practical experimentation, and iii. analyzing data had Cronbach's alpha values of the reliability of 0.789, 0.729 and 0.766 respectively. There is a relative strong inter correlations (Pearson) of the three FSWEx subscales. For example, the correlation coefficient (r) between planning experiments and carrying out practical experimentation is 0.567, planning experiments and analyzing data is 0.671 while carrying out practical experimentation and analyzing data is 0.619. Lastly, the scale correlates fairly well with its academic self-concept in Biology as a covariate. Damerau (2012) correlated FSWEx sub scales with the self-concept scale in Biology developed by Englin (2004) and found that the coefficient (r) of planning for experiments was 0.336, the coefficient (r) for carrying out experiments was 0.550 while that of analyzing data was 0.554.

6.2.2.4 Genetic test for measuring conceptual understanding of Genetics

To assess genetics knowledge as a covariate, a multiple-choice (single-select) item test containing 25 items was developed (see Appendix IV). A number of sources were reviewed for possible test items, including the example questions provided by the College Board's Advanced Placement Biology Exam, the SAT II Biology Exam, and the Biological Science Curriculum. Suitable items were ultimately included in a pool of questions. The test measures five (05) subtopics in Genetics as listed in the Tanzania Biology syllabus for the advanced level students which include i. hereditary materials (DNA/RNA), ii. genetic coding and protein synthesis, iii. Mendelian and Non-mendelian inheritance, v. sex-linked inheritance and pedigree analysis, and v. gene and chromosomal mutation. The test was reviewed by the supervisor of this study who is a professor of zoology and didactics of Biology to assure its content validity. A panel of three science educators further determined the content validity and clarity of each item on the test. The science teachers also analyzed the relatedness of the test items to the instructional objectives. They confirmed that the content validity of the instrument was appropriate for the participants. However, psychometric validation of this conceptual test was beyond the scope of this study. For scoring purposes, each multiple-choice item was given a numeric value of 1 if the response was correct or 0 if the response was incorrect. Therefore, scores ranged from 0 to 25.

6.2.2.5 The rationale of using the topic of genetics as a case study

As it has already been mentioned in section 1.4 of this report, genetics is one of the central topics addressed by the competence-based curriculum of 2005 in Tanzania for the advanced level Biology students. Genetics was taken as a case study because the topic is considered one of the most important and difficult topics in the school science curriculum (Tsui & Treagust, 2004). A number of reasons as why genetics concepts are difficult for students to learn have been reported by both teachers and researchers. For example, Pinar & Ceren (2008) indicated that these difficulties originate mainly from the domain-specific vocabulary and terminology, the mathematical content of Mendelian genetics, the cytological processes, the complex nature of genetics, and the abstract nature of the subject matter. According to Lewis & Wood-Robinson (2000), various genetics concepts depend on imaginary (theoretical) ideas constructed in abstract hypotheticodeductive conceptual systems. Therefore, a sound understanding of theoretical genetics concepts requires learners to reason hypothetico- deductively. Likewise, Banet & Ayuso (2000) argued that meaningful understanding of genetics is difficult and requires a certain level of abstract thought. Tsui & Treagust (2010) stressed the importance of having contemporary knowledge on DNA, genes, and their relations to human affairs on making informed decisions about ethically and socially controversial issues. Researchers in science education have consistently criticized the traditional teaching approach and suggested the development of more effective alternatives such as the inquiry-based approach.

6.2.3 Participants in the quasi experimental study

The participants of the study were 263 advanced level Biology students (age range 19-20) from selected secondary schools in Morogoro Tanzania. Three schools namely Kilakala (145 students), Alfagerms (87 students) and Bigwa sisters (31 students) were involved in the study. Activities that used inquiry, hands-on models and problem-solving were targeted for form five students while a lecture method was employed to teach form six students. This is because of the fact that Form six students didn't have much time for inquiry activities. These are finalist students and always busy for the preparation of their final national examination. The students, divided into an experimental (169 students) and a control group (94 students), attended a Biology course that involved themes on modern genetics and Mendelian inheritance topics. As summarized in table 6.1 below, the number of female students involved was 200 (130 in inquiry classes and 70 in conventional lecture method) while there were 63 male students 24 being in conventional lecture approach and 39 were involved in inquiry classes. The emphasis was on the understanding of the nature, function and correlations between the basic genetic concepts (e.g. DNA, genes, chromosomes, and meiosis) and the phenomenon of Mendelian inheritance protein synthesis and Mutation. None of the participants had been taught genetics at higher levels in the past.

Table 6.1: Distribution of students by type of instruction and sex in each school

			Sex		Total
			Female	Male	
Kilakala sec school	Instruction	Conventional approach	49		49
		Inquiry based method	96		96
	Total		**145**		**145**
Alfagerms	Instruction	Conventional approach	7	24	31
		Inquiry based method	17	39	56
	Total		**24**	**63**	**87**
Bigwa Sisters	Instruction	Conventional approach	14		14
		Inquiry based method	17		17
	Total		**31**		**31**
	Grand total		**200**	**63**	**263**

Source: Research survey (2014)

6.2.4 Controlling teacher factors/variables

Review of research literature has led to the conclusion that it is the teacher, more than the material, the method, or any other variable, that makes the greatest difference in children's educational achievement (Wright et al. 1997; Hattie, 2009). Teacher factors such as self- efficacy, interest, attitude, qualification, motivation, experience, knowledge, skills, teaching competence can not be ignored as can have profound impacts on various students' learning outcomes (Wang et al. 1993). At the heart of this line of inquiry is the core belief that teachers make a difference. For instance, teachers who demonstrate patience, knowledge of intervention techniques, an ability to collaborate with an interdisciplinary team, and a positive attitude towards children can have a positive impact on student learning success and the vice versa is true. Hence the influence of teacher variables in student learning outcomes cannot be ignored. In order to control the influence of teacher variables in this study, both the control and experimental groups were taught themes of Genetics by the researcher only who is also a Biology teacher. The researcher taught Genetics to the control group using conventional lecture method and the experimental group using inquiry-based approach. This means that differences in students' performance if there are any, can directly be attributed to the effectiveness of the method of teaching rather than the influence of teacher variables.

6.2.5. Implementation of genetics to the control and experimental classes

6.2.5.1 Implementation of genetics to the control group

Conventional lecture method was employed to teach themes within genetics to form six student classes in the selected schools. Lecture notes and discussion questions were prepared in advance before the actual class session. Three different textbooks prescribed by the Tanzania Biology syllabus and proved adequate to provide the essential factual basis for the course and were used in the construction of student's notes and discussion questions. They included Biological Sciences by Taylor D. J et al. (2008), Understanding Biology for Advanced Level by Glenn Toole and Susan Toole and Advanced Biology Principles and Applications by D.J Mackean. Each subunit met a total of 240 min/week (either 80 min on Monday/Wednesday/Friday or 120 min on

Tuesday/Thursday) plus a 50-min recitation each week for a total of 8 weeks. Topics discussed included i. hereditary materials (DNA/RNA), ii. genetic coding and protein synthesis, iii. Mendelian genetics iv. Non-mendelian inheritance and pedigree analysis, v. Gene and chromosomal mutation vi. meiotic and mitotic chromosome behavior, including recombination, mapping, and chromosome aberrations. Posttest scores of students were reported back to their respective Biology teachers at the end of intervention so that remedial measures could be taken for those who didn't perform well in this genetics test. Student marks were also supposed to be included in their total coursework results.

6.2.5.2 Implementation of genetics to the experimental group

Activities that used inquiry, hands-on models and problem-solving were targeted for form five students in the selected schools. As it has already been stated (see section 1.93 and 1.94) 5E instructional model (Bybee et al. 2006) and the constructivism formed the framework in teaching the experimental group. The role of the researcher in the experimental group was to promote discussion, active learning, and reflection and provide modeling, coaching, and scaffolding to students when required. As suggested by constructivists, the teacher (the researcher) acted as a facilitator rather the custodian of knowledge. In inquiry classes, many hours were dedicated to building new activities/models, and other activities. Throughout the teaching by inquiry, Biology students were working in small groups where they were encouraged to explore problems, formulate hypotheses, designing micro experiments share their ideas with their classmates, discuss their observations and interpret findings of the experiments or hands-on activity carried out. For example, students investigated some inherited and acquired human traits that are easy to observe in a classroom. Working in groups of four, students took a personal inventory of their traits (i.e. dimples, widow's peak, pierced ears, etc) and compare their traits to the rest of the class. In addition to introducing basic genetic terminology, this activity introduced the concepts such as the relationship between molecular differences in the DNA and observed physical traits and the difference between inherited and acquired traits. Students also had the opportunity to practice inquiry skills, make data tables, and analyze graphs. The study began with a pre-test assessment to gather information about the students' prior knowledge of genetics as the main objective.

The students in the inquiry experimental group worked in groups of four. The school Biology book was not used at all and the role of the teacher was reduced to that of a coordinator and facilitator of the students' work. The students' main learning aid was a set of worksheets which was collected from different sources mainly websites (see table 6.2 below) prepared specifically for the teaching of the genetics. The worksheets complete with short articles as a source of new information, tables, diagrams, pictures, exercises, and guidelines for small investigations, facilitated the application of the inquiry approach. Several small changes had to be made as the teaching progressed to adjust to the specific needs of the students and to support their investigations. At the beginning of some lessons, students were presented with a scientific phenomenon or set of data and were asked to make observations and specify relevant research questions after selecting an appropriate problem for investigation. The experimental group underwent a total of sixteen inquiry-based lessons, of which two lessons on average were accomplished per week in eight weeks as shown in table 6.2 below. Table 6.2 summarizes the sequence of activities that were implemented in this unit after pre-tests.

Table 6.2: Sequence of activities. This table includes all activities addressed during the genetics unit and their category as a hands-on model, problem solving, or inquiry-based activity.

Day	**Activity**	**Hands-on Models (M) Problem Solving (PS) Inquiry (I)**
Day 1	Pre-test	BPST, SMQ-II, FSWEx and Genetics test
Week 1	Chromosomes structure, Mitosis and meiosis	Discussion on Chromosomes structure and functions
		Mitosis hands on activity
		Meiosis Model Activity
Week 2	DNA as a hereditary material	Extracting DNA from Your Cells
		DNA replication: A case discussion of a landmark paper by Meselson and Stahl
Week 3	RNA and Protein synthesis	Protein Synthesis Modeling activity
		A case discussion of protein synthesis questions
Week 4	Mendelian Genetics	A class discussion of Mendel's pea plants experiment
		Modeling monohybrid crosses activity
		Dihybrid Cross Activity (Busch Gardens, 2003) Problem Solving Activity
Week 5	Non Mendelian Genetics	Sponge Bob Incomplete Dominance Activity
		Using Blood Types to Solve a Mystery Class Activity
Week 6	Sex linked characters and pedigree analysis	Sex determination discussion activity
		Sex linked characteristics and the royal family pedigree problem solving activity
Week 7	Blood genetics and Lethal genes	Personal pedigree and analysis survey activity by Larry Flammer (2006)
		Blood Typing Murder Mystery Activity
Week 8	Gene and chromosomal mutations	DNA Mutations- Become a Genetic Counselor
		mutation inquiry activity questions
Final day	Post-test	BPST, SMQ-II, FSWEx and Genetics test

Nature of inheritance activity By Dr. Ingrid Waldron, Department of Biology, University of Pennsylvania, 2015

This lesson was designed by Dr. Ingrid Waldron of the department of Biology, at the University of Pennsylvania in 2015. This activity guided students to understand basic genetics concepts, including how genotype influences phenotype and how understanding meiosis and fertilization provides the basis for understanding inheritance. Students were also supposed to understand how genes influence our characteristics. The activity in this lesson included all of the basic concepts and introduces students to the Punnett square as a summary of how genes are transmitted from parents to offspring by the processes of meiosis and fertilization. Students learned that hereditary information is contained in genes, located in the chromosomes of each cell and that cells contain many thousands of different genes. One or many genes can determine an inherited trait of an individual, and a single gene can influence more than one trait. On the other hand, students were supposed to explain how DNA/ chromosomes are inherited from parent sex cells to offspring during sexual reproduction. Throughout, students responded to analysis and discussion questions to further develop their understanding of how genes influence our behavior.

Mitosis - how each new cell gets a complete set of genes, By Drs. Ingrid Waldron, Jennifer Doherty, R. Scott Poethig, and Lori Spindler, Department of Biology, University of Pennsylvania, 2015

This minds-on, hands-on activity helped students to understand how mitosis ensures that each new cell gets a complete set of genes. The instructional philosophy for this activity and follow-up activity on meiosis and fertilization was that student learns about mitosis in a most meaningful way. Furthermore, students learned how gene-carrying chromosomes move during mitosis understand how these processes result in transmission of genes from parents to offspring. To provide the background needed for this approach, this mitosis activity begun with an introduction to chromosomes, genes and alleles, and the effects of genes on phenotypic characteristics. Then, students learn about the basic process of mitosis and use model chromosomes to simulate mitosis. Students manipulate pipe-cleaner chromosomes on a template showing stages of mitosis with one pair of chromosomes until approved by the teacher. Students understood the verbs associated with mitosis process and gain a sense of how the cell must sequence the many steps involved in mitosis.

Meiosis Models Hands-on Model Activity

This is a free genetics activity for secondary school students found in http://www.nclark.net/ModelingMeiosislab. In this activity, papers were used to make chromosome models. In pairs, students manipulated three homologous pairs through the phases of meiosis. This occurred after an in-depth class discussion of gamete formation and the phases of meiosis. The students were to physically manipulate the models through meiosis one and meiosis two and draw the chromosomes at certain key phases during these processes. These drawings enabled students to see how chromosomes line up differently during meiosis one and two, demonstrating how gametes have different genetic combinations. At the end of the lesson, students learned that meiosis is a type of cell division that reduces the number of chromosomes in the parent cell by half and produces four gamete cells. During reproduction, when the sperm and egg unite to form a single cell, the number of chromosomes is restored in the offspring. At the end of the activity, the students answered discussion question to demonstrate whether or not they understood the entire process.

DNA structure and extraction activity by Drs. Ingrid Waldron, Jennifer Doherty, R. Scott Poethig, and Lori Spindler, Department of Biology, University of Pennsylvania, 2015

In this activity, students extract DNA from their cheek cells and relate the steps in the procedure to the characteristics of cells and biological molecules. Students learnt key concepts about DNA function during the intervals required for the extraction procedure. Student understanding of DNA structure, function and replication was further developed by additional analysis and discussion questions and hands-on modeling of DNA replication. The lesson provided an alternative activity in which students extracted DNA from strawberries and their cheek cells.

DNA Replication: A case discussion of a landmark paper by Meselson and Stahl (1958)

Before a cell divides, the genetic information must be copied and apportioned evenly into the daughter cells. This exercise couples a classic primary literature paper detailing the process of DNA replication with a set of questions designed to both guide students through the process of reading papers and delve deeply into the critical concept of replication. The paper by Meselson and Stahl (1958) is an ideal paper to introduce students to the art of reading papers and appreciating the beauty of science. This paper not only was a landmark experiment for the essential process it helped to define, but it was also recognized for its elegant simplicity. This resource is a facilitator's guide to help run a discussion session for advanced level Biology students.

Protein synthesis hands-on model activity by Ann Hoppe, S (2013)

In this hands-on activity, students learn how a gene provides the instructions for making a protein. The lesson was adopted from the dissertation project submitted by Ann Hoppe, S (2013) to the Michigan State University as a partial fulfillment of the requirements for the degree of Master of Science. Protein synthesis kits were constructed from craft foam and Velcro. The kits include a DNA backbone, mRNA backbone, nitrogen bases (nucleotides), tRNA, amino acids, and peptide bonds. The students constructed a nine base DNA strand and transcribed it into mRNA, which was used as a template for protein synthesis. During the activity, the students answered process questions and wrote down the DNA base sequence, mRNA codons, and amino acid sequence. After this activity was completed the students did it again in reverse to reinforce the processes. They started with three amino acids beginning with methionine, the start codon. Then, the students have to find the mRNA sequences and DNA that corresponded to the amino acid code. This activity was used to introduce students to transcription and translation or to reinforce student understanding.

Modeling Mendel´s pea plant experiment

This modeling activity allows learners to discover for themselves what Mendel uncovered in his famous pea experiments. By modeling Mendel's pea experiments, learners formed their own explanations for the result of crossing a true-breeding round pea plant with a true-breeding wrinkled pea plant (the F1 generation) and for the results of allowing an F1 pea plant to self-pollinate (the F2 generation). They then compared their explanations to Mendel's own conclusions. At the beginning of the activity, students were asked questions such as: What did Mendel use as a subject of his experimentation? How did he predict possible offspring results? This is an excellent introduction to Mendelian genetics which generated discussion and stimulated interest in Mendel's principles. Learners were encouraged to use the same observation and critical thinking skills that Mendel used.

Sponge Bob genetics problem solving activity (Trimpe, 2003)

In this unit, problem-solving activities were implemented that used a familiar test subject, "Sponge Bob". The concepts addressed in this activity were dominant and recessive alleles, homozygous and heterozygous pairs; and test crosses/punnet squares were conducted to determine expected outcomes. The students conducted the Sponge Bob crosses and calculate the genotypic and phenotypic ratio and percentages of offspring with certain alleles. The students were given different scenarios using Sponge Bob and his friends as examples in order to give them a variety of problems in repetition without it becoming tedious. The students were more interested and engaged in these problems than when data from pea crosses were used in the past. In another part of this activity, students had to determine the genotypic and phenotypic ratios of test crosses to determine how incomplete dominance is different from typical dominant and recessive inheritance.

Modeling monohybrid crosses problem solving activity

This lesson focuses on the first Mendelian law of inheritance and was adopted from *www.central.rcs.k12.tn.us/teachers/gullettp/documents/modelingmonhybridcrosses.pdf* and it is the copyright by Holt, Rinehart and Winston (2006). In this lesson, students predicted the genotypic and phenotypic ratios of offspring resulting from the random pairing of gametes and then calculated the genotypic ratio and phenotypic ratio among the offspring of a monohybrid cross. A monohybrid cross is a cross that involves one pair of contrasting traits. Different versions of a gene are called alleles. When two different alleles are present and one is expressed completely and the other is not, the expressed allele is dominant and the unexpressed allele is recessive. In this practical simulation some students set up crossings using either small squares (2 cm x 2 cm) of red colored paper and white colored paper OR beads (red and white) to study allele combinations for a monohybrid cross.

Dihybrid crosses problem solving activity by Busch Gardens (2003)

This activity from the SeaWorld website focused on five different crosses using all kinds of unique animals as test subjects. The class was divided into groups of four and one case study funsheet to each group was distributed. Hence each group had a different case study funsheet to calculate. Students worked in groups and completed the dihybrid crosses. They answered questions about the cross after completing the punnet squares. The students had to calculate phenotypic and genotypic ratios of the offspring. They then discussed the Mendelian law of independent assortment of gametes during meiosis and fertilization.

Patterns of inheritance activity by Cody Alley and Jake Skinner (2013)

The purpose of this activity was to delineate the different types of gene expression and how alleles interact with each other. Students worked on exercises to help solidify the concepts of dominance, codominance, incomplete dominance, etc. The exercises involved mixing of genes and reading pedigrees. The Students were supposed to explain what a multiple allele trait is, identify blood type as a multiple allele trait and create Punnett Squares to interpret crosses of multiple allele traits. The student investigated and understood common mechanisms of inheritance. The key concepts included prediction of inheritance of traits based on the Mendelian laws of heredity.

Blood typing murder mystery problem solving activity By Drs. Jennifer Doherty and Ingrid Waldron, Dept Biology, Univ Pennsylvania, © 2013.

In this activity, students learn the genetics and immunobiology of the ABO blood type system and use simple chemicals and logical reasoning to solve a murder mystery. Students answered some basic questions about the genetics of blood type, multiple alleles, and co-dominance, and then completed a problem-solving activity. The students read a mystery about a millionaire who died and a man who claimed to be his kin to collect his inheritance. The students were given the blood type of the millionaire and the alleged son. The students used the blood types to prove that the man was or was not his offspring by completing punnett squares. The purpose of this lesson was to explore multiple allele traits. In another assignment, students were given the scenario where they were supposed to determine whether two babies were switched in the hospital or not. Students learned how to perform Punnett Square crosses for blood type, a multiple allele trait. They then applied this knowledge to mystery scenario and determine the genotypes for each individual involved, and use at least two Punnett Squares as evidence.

Sex determination problem solving activity

This lesson activity was developed by Peter Berry of Middletown High School and Scott Holmes of Wesleyan University in 2009 and found in http://www.ashg.org/cgibin/gena/glesson.pl?s=LSN&t=1&l=1&c=0 . This lesson explores sex determination in mammals. Students first share their conceptions about the genetics of sex determination, then using normal and abnormal human karyotypes, students hone in on their hypothesis. The case of XY females was used to explain the presence of the SRY gene on the Y chromosome, which must produce an active protein in order for a male to develop. The lesson ends by exploring other sex-determination systems in animals and their similarities and differences. At the end of the activity the students answered analytical questions to demonstrate whether or not they understood the entire process of sex determination in human beings.

Sex - linked characteristics and the royal family pedigree problem solving by Lauren Woodside, Kentwood High School

This inquiry genetics lesson was designed by Lauren Woodside of Kentwood High School to her 10th grade Genetics students. The main objective of this activity was to enable students describe the role of sex chromosomes in reproduction and understand the role of probability in the study of genetics and inheritance. Students learned the names, characteristics, and causes of several different sex-linked disorders, including muscular dystrophy. Students created, described and predicted genotypes according to genetic pedigrees. This learning cycle goes a step beyond basic heredity as the student follows colour blindness genes through a family tree. The student viewed how sex-linked traits in heredity work and how the desired genes are usually dominant over less desirable genes. The students were introduced to the Romanov family – the Tsar, Tsarina, and their five children. Students learned that the youngest son Alexei inherited the sex-linked genetic disorder hemophilia and see how the disorder was passed through members of the royal family. Through computer links, students learned more about the inheritance of hemophilia. Some guiding questions were,

i. Which individuals are carriers of hemophilia in this family?
ii. Which individuals have hemophilia in the family?
iii. What trends or patterns do you see in who gets this genetic disease?
iv. Why might you see the patterns you see?
v. What differences do you see between female chromosomes and male chromosomes?
vi. Which chromosome do you think the gene for hemophilia is on?

Personal pedigree and analysis survey activity by Larry Flammer (2006)

This adopted lesson was designed by the Evolution & the Nature of Science Institutes (ENSI) *www.indiana.edu/~ensiweb* in 2006. The lesson is a most engaging activity in which students connect their own family to the principles of Mendelian genetics. Students were asked to survey their family (or a neighbor or friend's family), looking for a trait that shows different phenotypes. They diagram the family using pedigree symbols, add the phenotypes where known, then look for the discriminating pattern that tells which phenotype is dominant. (Students learn that it's not necessarily the most common one). From there, genotypes were inserted for each person and even possible phenotypes/genotypes for persons whose phenotypes were unknown indicated. The lesson intended to recognize the pattern in family trees that tells us which phenotypes are dominant and which are recessive.

DNA mutation inquiry-based activity

This lesson has been published by NGSS life science and accessed through hhttp://www.ngsslifescience.com/science.php/biology/lessonplans/C408/. In this lesson, students learned the effect of DNA mutations on protein formation and phenotype. This activity is inquiry based where students mimic chromosome mutations by cutting and pasting paper chromosomes. The students converted a DNA sequence to an amino acid sequence and use color-by-number pictures to show the difference between an original and mutated sequence. Lesson activities were structured to support student comprehension of the roles of proteins, how the genetic code is used, causes of mutations and what mutation can mean to an organism in its environment. Through comparisons with other students in the class, the students learned that not all mutations result in a change, while some may cause a great deal of change in a gene (and therefore the protein and/or phenotype). Students then used this activity to defend a claim that inheritable variations can be caused by mutations.

6.2.6 Inquiry based teaching intervention and student's motivation

This inquiry-based teaching intervention of genetics was not designed specifically to contribute to students' attitudes towards science. Yet, it was assumed that it might have a positive effect on the students' attitudes as they were free to choose what to investigate, use their own method to follow and conduct the investigation on their own pace. In this respect, the students' attitudes towards science were scored prior to and after the teaching intervention. Hence another purpose of this quasi-experimental study was to examine students' motivation towards science process skills learning outcome after implementing inquiry instruction on advanced level Biology students in Morogoro. Science Motivation Questionnaire II by Glynn et al. (2011) and FSWEx Self-concept scale by Damerau (2012) were used as data collection tools for students' motivation. Inquiry-based learning emphasizes learning by doing and mirrors the work of scientists as they actively discover knowledge. The study answered a specific question 'what are the differences in students' motivation towards science process skills after they experienced inquiry and traditional genetics instructions?

6.2.7 Administration of tests

The tests were administered at the beginning (pretests) and at the end of genetics course intervention (posttests) to ensure that all subjects have undergone approximately the same science program. To minimize disruption of teaching in the classes involved and to avoid fatigue as a result of taking two tests successively, the two tests (BPST- 60 minutes and genetics test- 35 minutes) were administered on two different days within the same week. The tests were administered in the same week in order to minimize the effect of learning that would have occurred in between the administration of the tests. There were no data losses because schools involved were boarding schools at which all of the students live during the part of the year that they go to lessons. So it was easy to control their class attendance. Additionally, students completed the Science Motivation Questionnaire-II and the FSWEx self-concept scale by Damerau (2012) together with the genetics test at the sometime because they didn't require much time.

6.2.8 Data analysis plan

Both data collection tools in this particular study (science process skills test, genetics test, motivation Likert scale and self-concept scale) provided quantitative data. These data were analyzed using SPSS version 21.0. The groups were given both the pre-test and the post-tests (science process skills, Genetics, self-concept and motivation tests). The overall pretest and posttest scores from Biology process skills test (BPST) and genetics test were calculated for each student in terms of the percentage of correct responses. These scores were analyzed in several ways. First, a general linear model was used to determine, whether there are statistical differences between the experimental and control groups in terms of their performance in the science process skills, genetics, and motivation tests. A repeated measure analysis of variance was used to analyze the effect of time. It is the statistical measure used to examine multiple observations of a scale overtime and/ or under different conditions (Schindler, 2014; Green et al. 2000). In this study repeated measures analysis of variance (ANOVA) was conducted to test for between-group differences overtime. Repeated measures ANOVA for between - group differences is entitled Multivariate analysis and a repeated measures ANOVA for within - group differences is entitled test of within - subjects effect. The measurement of time consists of time elapsed over 08 weeks of each aspect of study with measurement at pre-test (week 01) and post-test (week 08).

In a within-subjects test, the sphericity assumption can be a problem. This assumption states that correlations between all pairs of measurements are roughly the same. This means that measurements of several different times are not particularly robust to this assumption (Schindler, 2014; Sherry, 1997). This problem is common with repeated measures over time. However, the Greenhouse-Geisser correlation will fix the sphericity problem (Schindler, 2014; Sherry, 1997). Multivariate analysis of variance was conducted to measure between group effects overtime. All students in the control and experimental groups present during pretest also participated in the posttest (169 experimental group and 94 in the control groups). In the test of within-subjects, the within-subject factor was time with two levels (Pretest week 01 and posttest week 08) and the dependent variables are the scores in each of the instrument at these levels (pretest and posttest).

Secondly, t-tests for paired samples were performed on the pre- to posttest difference scores (pretest scores subtracted from the posttest scores) for all participating students to test for statistically significant differences between pretest and posttest scores. A t-test was used to test differences between two means because of its superior quality in detecting differences between two means (Borg & Gall, 1996). All tests of significance were tested at a significance level of 0.05. Thirdly, bivariate analysis using Pearson correlation were used to determine the influence of motivation, self-concept, and genetics conceptual understanding of students' on their performance in science process skills. Fourthly, descriptive analysis of frequencies, percentages, means, and standard deviations was used to organize and analyze data from these instruments.

6.3 RESULTS AND DISCUSSION OF THE QUASI-EXPERIMENTAL INTERVENTION STUDY

6.3.1 Introduction

The inquiry-based approach is an inductive constructivism pedagogy that places students' questions, ideas, and observations at the center of learning experience (Rolf et al. 2010). It emphasizes the need for learners to construct meaning out of the lesson. A conventional method, on the other hand, is a direct instruction model for teaching that emphasizes well-developed and carefully planned lessons designed around small learning increments and clearly defined and prescribed teaching tasks. In this model of teaching, teachers provide instructions, ask leading questions and dictate relevant information to the topic of discussion(Chung, 2004). In contrast, an inquiry-based approach challenges students to learn by observing, experimenting and asking questions. This study aimed at comparing an inquiry-based approach vs a conventional method in the development of students' science process skills, understanding of contents and development of motivation towards science. As it has been written earlier, the current Tanzania science education reforms (URT, 2005 & 2010) has placed an extensive emphasis on scientific inquiry in which students are expected to master a set of inquiry-related skills and develop understandings about inquiry as opposed to mastery of science contents only.

This chapter presents and discusses key findings of the quasi-experimental study to compare the effectiveness of inquiry-based teaching approach and conventional method on students' scientific process skills development, conceptual understanding of Biology contents and motivation towards science process skills. Students completed the same data collection instruments before and after intervention so that changes in their conceptual understanding, scientific process skills and motivation towards science could be spotted. Although the revised competence based curriculum of Tanzania (MoEC, 2005) emphasizes the use of learner-centered activity-based pedagogy (inquiry-based approach) especially in sciences, as opposed to conventional methods. However, not much work has been done in Tanzania to assess the effectiveness of this teaching method. This void is the essence of the current quasi- experimental study. Data collection tools used included researchers' constructed genetics test, a validated science process skills test (BPST), Science Motivation Questionnaire II (SMQ-II) by Glynn et al. (2011) and FSWEx questionnaire by Damerau (2012). As it has been described in the introduction part of this chapter (section 8.2) this stage was an intervention phase and it intended to answer statistically the following major questions:

i. Is there statistically significant difference in science process skills achievement between students exposed to inquiry-based approach (IBA) and those exposed to a conventional or traditional method (TM)?
ii. Is there statistically significant difference in the achievement of specific science process skills understudy (hypothesis formulation, controlling variables, designing experiments, analyzing data and defining operationally) between students exposed to inquiry-based approach (IBA) and those exposed to the conventional/ traditional method (TM)?
iii. Is there statistically significant difference in the conceptual understanding of genetics contents between students exposed to inquiry-based teaching (IBA) approach and those exposed to traditional method (TM)?
iv. Is there statistically significant difference in motivation towards science (intrinsic motivation, self-efficacy, and self-concept) between students exposed to inquiry-based teaching (IBT) approach and those exposed to traditional methods (TM)?
v. How is the students achievement in science process skills correlate with their i. conceptual understanding, ii. intrinsic motivation, iii. self-efficacy, iv. self-determination, and v. self-concept towards science?

The findings are presented in four main sections; section 6.3.2 presents findings on the development of students science process skills, section 6.3.3 the findings on students' conceptual understanding of genetics, section 6.3.4 presents findings on students' development of motivation towards science process skills and lastly section 6.3.5 presents the findings on the correlation between students achievement in science process skills with i. conceptual understanding, ii. intrinsic motivation, iii. self-efficacy, iv. self-determination, and v. self-concept towards science.

6.3.2 RESULTS WITH RESPECT TO STUDENTS' LEVEL OF SCIENCE PROCESS SKILLS (FINDINGS FROM BPST)

6.3.2.1. Science process skills pretest results from descriptive and t-test

The major aim of this quasi-experimental study was to find out whether there is a statistically significant difference in science process skills achievement between students exposed to inquiry-based teaching (IBA) approach and those exposed to traditional method (TM). The study involved 94 (35.7%) control group students who were taught themes in genetics using the conventional method and 169(64.3%) experimental group students who were taught genetics using inquiry-based approach (IBA). An SPSS two-tailed independent samples t-test was conducted to compare pretest scores of experimental (IBA) and control (TM) classes on science process skill test (BPST) before the actual intervention. The pre-test was administered to the experimental group and the control group in order to determine whether the two groups were similar in terms of their level of science process skills before teaching intervention. Because the two groups were composed of advanced level students who are taking Biology and the fact that they are undergoing the same curricular materials, the study hypothesized that the two groups would not significantly differ in terms of their level of science process skills. Using independent-samples t-test and descriptive statistics this hypothesis and was tested (see table 6.3a and b).

As it has been summarized in table 6.3 (a) and in the figure 6.1 section 6.3.2.3, the mean of pretest scores in BPST for the students in the control group was 15.2 out of 35 (one mark for each of the 35 items) while the mean of scores for the students in the experimental group was 15.4. The standard deviation (a spread of individual scores around their respective means) was 2.84 for the control group and 2.44 for the experimental group. This means that before genetics course intervention, the variability of the control group (2.84) was more than that of the experimental group (2.44) as shown by the coefficient of variation. This implies that the experimental group was more homogenous in terms of science process skills level than the control group before the intervention.

Table 6.3(a): *Group statistics for BPST pretest scores based on the type of instruction they received (n=263)*

	Grade level of the students	*N*	*Mean*	*Std. Deviation*
Grade level of the students	Control group	94	15.2	2.84
	Experimental group	169	15.4	2.44

Source: Field data (2015).

To verify that the two groups were matched on pretest scores in the science process skills test and provide justification for interpreting gain scores for the sample, independent samples t-tests were performed comparing the inquiry and control group on pretest measures. However, no statistically significant differences in the level of science process skills were found among students of the control group and experimental group when their pretest mean scores were subjected to computer SPSS independent samples t-test. As shown in table 6.3 (b) below, the results of pretest scores of the IBA group (M=15.4, SD= 2.24) and that of TM classes (M=15.2, SD= 2.84); found t (261) =-1.403, p = 0.224, hence p > 0.05.The earlier hypothesis that the two groups do not significantly differ in terms of their science process skills knowledge level was accepted.

Table 6.3(b): *Independent samples t-test for BPST pretest scores based on the type of instruction they received (n=263)*

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Students level of process skills	Equal variances assumed	4.225	0.041	-1.403	261	0.194	0.37324	-0.41546	-1.04311	0.21219
	Equal variances not assumed			-1.220	158.4	0.224	0.37324	-0.41546	-1.08791	0.25699

Source: Field data (2015).

These results from table 6.3 (a and b) above suggest that the knowledge levels of science process skills of students both in the control and experimental groups were comparable prior to genetics teaching intervention. According to Reinhart & Rallis (1994) in quasi-experimental pretest-posttest studies, if groups differ at the onset of the study, any differences that occur in test scores at the conclusion will be difficult to interpret. The experimental and control groups of Morogoro Biology students, in this case, were therefore regarded suitable for this kind of a comparative study.

6.3.2.2 General linear model pretest- posttest results comparison for control and experimental groups

The study was also interested in examining the within and between-group differences with respect to the development of students' science process skills over time as a result of the intervention. The best method of analyzing quasi-experimental data is to view the pretest and posttest as a repeated measures/ split-plot design or as a profile of two measurements for each subject (Green et al. 2000). According to Field (2006), repeated measures can be used to observe both the within-person (or within-subject effects and the between-persons (or between-subjects) effects. A within-person (or within-subject) effects represent the variability of a particular value for individuals in a sample. Between-persons (or between-subjects) effects, by contrast, examine differences between individuals. According to Shuttleworth (2009) between-subjects is an experiment that has two or more groups of subjects each being tested by a different testing factor simultaneously. However, quasi-experimental data are commonly examined in repeated measures analysis. A repeated measures analysis is a measure of how much an individual in the sample tends to change (or vary) over time. In other words, it is the mean of the change for the average individual case in the sample and it is observed in one and only one treatment combination (Martin, 1996).

In this study, the SPSS general linear model for repeated measures was conducted to test the effectiveness of both the conventional and inquiry methods for within- and between- groups differences in science process skills development over time. Repeated measures ANOVA for between - group differences is entitled "the effect of time on groups" and a repeated measures ANOVA for within - group differences is entitled a test of interactions effect on groups (Schindler, 2014; Green et al. 2000).

Repeated measures analysis of variance (ANOVAs) were conducted on science process skills test (BPST) scores to compare groups' scores over the two testing occasions to test for between and within-group differences over time. The measurement of time consists of time elapsed over 08 weeks for each aspect of study with measurement at pretest (week1) and post-test (week 08). In the test of within-subjects, the within-subject factor was time with two levels (pretest in week 01 and posttest week in 08) and the dependent variables is the BPST scores at the pretest and posttest levels. On the other hand, in the test of between - groups difference, the factor was the two groups (control n= 94 and experimental group n= 169) overtime (pretest week and posttest week 8) and the dependent variable was student scores in the Biology process skills test (BPST). The findings from SPSS general linear model for repeated measure (within and between groups) are presented in section 6.3.2.2.1 and 6.3.2.2.2.

6.3.2.2.1 ANOVA findings for within group (test of within - subjects effect)

A within subjects ANOVA was performed on science process skills test (BPST) scores to compare groups score over the two testing occasions. As it has already been stated in section 6.3.2.2 above, in the test of within-subjects, the within-subject factor was time with two levels (pretest and posttest) and the dependent variable was student scores in the BPST (pretest and posttest). The intention was to test the significance of a mean gain score of the experimental and control group in the achievement in science process skills. A Repeated measures analysis of variance is the statistical measure used to examine multiple observations of scale over time and/ or under different conditions (Schindler, 2014; Green et al. 2000). The ANOVA tested the null hypothesis which stated that there is no statistically significant within groups (control and experimental) in the acquisition of science process skills (control and experimental) for two time periods (pretest and post test). As it has already been stated in section 6.3.2.1 above, the study involved 94 (35.7%) control group students who were taught genetics using the conventional method and 169(64.3%) experimental group students who were taught using inquiry-based approach (IBA).

Table 6.4 (a) summarizes the findings of SPSS general linear model with repeated measure for pretest and posttest scores within groups (experimental and control groups). SPSS computation of the general linear model with repeated measure within groups found $F(1,261) = 471.081$, $p < 0.001$, eta squared =0.643. Hence a significant main effect was noted for the time, $F(1, 261) = 471$, $p < 0.001$, which means regardless of the method of teaching there was a significant within groups effect on the development of science process skills as a result of the methods of teaching. The null hypothesis which stated that there is no statistically significant within-group effect between the control group and the experimental group over two testing occasions (pretest and post test) with regard to students' science process skills development was rejected. Statistical significant time effects were noted at alpha =0.05 level.

Table 6.4(a): ***Within-subjects effects on science process skills for two time periods (control group n= 94 & experimental group n= 169)***

Tests of Within-Subjects Effects							
Measure							
Source		Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Test scores (BPST´)	Sphericity Assumed	2279.038	1	2279.038	471.081	0.000	0.643

Source: Field data (2015)

According to table 6.4(a), the eta square value was acquired as 0.643. This effect size value shows that the effect magnitude is large and that almost 64.3%. This further implies that 64.3% of the change observed in the dependent variable resulted from the application of the treatments (methods of teaching). This means that both teaching methods (inquiry-based approach and conventional method) used in this study create a statistically significant difference in Morogoro Biology students' science process skills disposition scores. However, results of several studies (Rissing et al. 2009; Marx et al. 2006) have shown that student' scientific process skills can be developed by using inquiry or investigative approach of teaching and learning science that gives them opportunities to practice these skills than the traditional method.

6.3.2.2.2 ANOVA for between - group differences on science process skills (Test of between - subject effects)

The between-subjects effects determine if respondents differ on the dependent variable, depending on their group or depending on their score on a particular measure (Shuttleworth, 2009). A comparison of the groups tells about the effects of the treatments. The variability of scores within each group reflects individual differences as a result of treatment. A repeated measures ANOVA was conducted with the factor being the two groups (control n= 94 and experimental group n= 169) overtime (pretest week and posttest week 8) and the dependent variable being student scores in the Biology process skills test (BPST). The results are presented in table 6.4 (b). According to table 6.4(b) below, the test of between - subject effects found F (1, 261) =0.471.081, p < 0.157 which means that the linear model accepted the null hypothesis. This further implies that the between-group interaction effects (method * groups* time) was not significant. Hence the null hypothesis was accepted at alpha = 0.05 level. The null hypothesis stated that "there is no statistical significant between students exposed to the inquiry-based approach (IBA) and the traditional method (TM) in their development of science process skills over time. The within-subject test indicates that the interaction of time and the group was not significant. Taking into account the findings from within-group effects, this means that there were significant gains over time and but there was no statistically significant differential improvement among groups over time. The main findings showed that both methods had an impact on the development of scientific process skills to Morogoro students.

Table 6.4(b): Between-subjects on science process skills for two time periods (control group n= 94 & experimental group n= 169)

Tests of within-subjects effects							
Measure							
Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
BPST *Type of instruction	Sphericity Assumed	9.760	1	9.760	2.018	0.157	0.008

Source: Field data (2015).

The findings in table 6.4 (b) implies further that regardless of the teaching method, there was an improvement of science process skills to Morogoro Biology students both in the control and experimental groups with time. These results, however, do not support anecdotal claims that the inquiry-based method of teaching is more effective than the traditional lecture method of teaching in science process skills development. These findings led the researcher to conclude that there is no a single best or effective method of teaching in each context. Effective teaching method according to Seldin (1999) is any approach which produces beneficial and purposeful student learning through the use of appropriate procedures. For example, in this case, the study indicated interaction effect between time and treatment groups meaning that the experimental and control groups but had no significant differential improvements over time.

6.3.2.3 POSTTEST RESULTS ON STUDENTS SCIENCE PROCESS SKILLS

6.3.2.3.1 Comparing the general performance of control and experimental group in BPST

To determine statistically if there were a significant difference in students' science process skills achievement between those exposed to the inquiry-based teaching of genetics and those exposed to the traditional method, an analysis of BPST posttest mean scores was carried out. Two independent samples t-test was conducted to follow up the significant interaction and assess differences among teaching method groups at the end of intervention period. The two groups (control and experimental) were firstly given the pre-test and then and intervention of 08 weeks before completing the same post-tests. The testing effects and influence of teacher variables across all the groups were nullified so that the post-tests of each of the experimental groups could be compared with that of the control group to detect the effects of treatment/ intervention (see section 6.2.4).

The mean scores and standard deviations of two groups are shown in table 6.5 (a). The results of mean scores between the control and experimental groups on BPST have been represented also in a bar graph in Figure 6.2. The mean of students score in the experimental group was 19.9 out of 35, while the mean of the control group 19.1 out of 35 items. The spread (standard deviation) of individual scores around their respective means changed from 2.84 to 1.82 for the control group and from 2.44 to 1.97 for the experimental group. Contrary to pretest results, the variability the experimental group (1.97) was more than that of the control group (1.82) as shown by the coefficient of variation in the table (6.5 a) below. Hence the experimental group, in this case, was found to be a bit more variable than the control group implying that the control group was more homogenous than the experimental group after intervention (posttest).

Table. 6.5(a) Group statistics for BPST posttest scores based on the type of instruction they received (n=263)

	Grade level of the students	*N*	*Mean*	*Std. Deviation*
Grade level of the students	Inquiry based approach	169	19.9	1.822
	Conventional method	94	19.1	1.973

Source: Field data (2015).

To establish whether the difference in mean scores between the control group and experimental group were statistically significantly or not, an SPSS independent samples t-test analysis was carried out. The results from independent samples t-test of mean scores are shown in table 6.5 (b). According to the table (6.5 b), a statistical significant difference was found on students' posttest scores based on the type of instruction they received when the null hypothesis (Ho1) was subjected to computer SPSS independent samples t-test. An analysis of independent samples t-test based on the type of instruction students received at (α) =0.05 produced a p of 0.047 and a t-value of 0.633, hence rejecting the null hypothesis (Ho1). The null hypothesis stated that "there is no statistically significant difference in students' science process skills achievement between those exposed to inquiry-based teaching (IBA) and those exposed to traditional method (TM)". It means that there was statistically significant difference in students' science process skills achievement between those exposed to inquiry-based (IBA) and those exposed to traditional method (TM) in favor of the experimental group. The null hypothesis was rejected at 0.05 alpha levels. Tables 6.5 (b) below summarize the independent samples t-test of scores based on students grade level.

Table 6.5(b): Independent samples t-test for BPST posttest scores based on type of instruction (n=94 control group & n=169 experimental group)

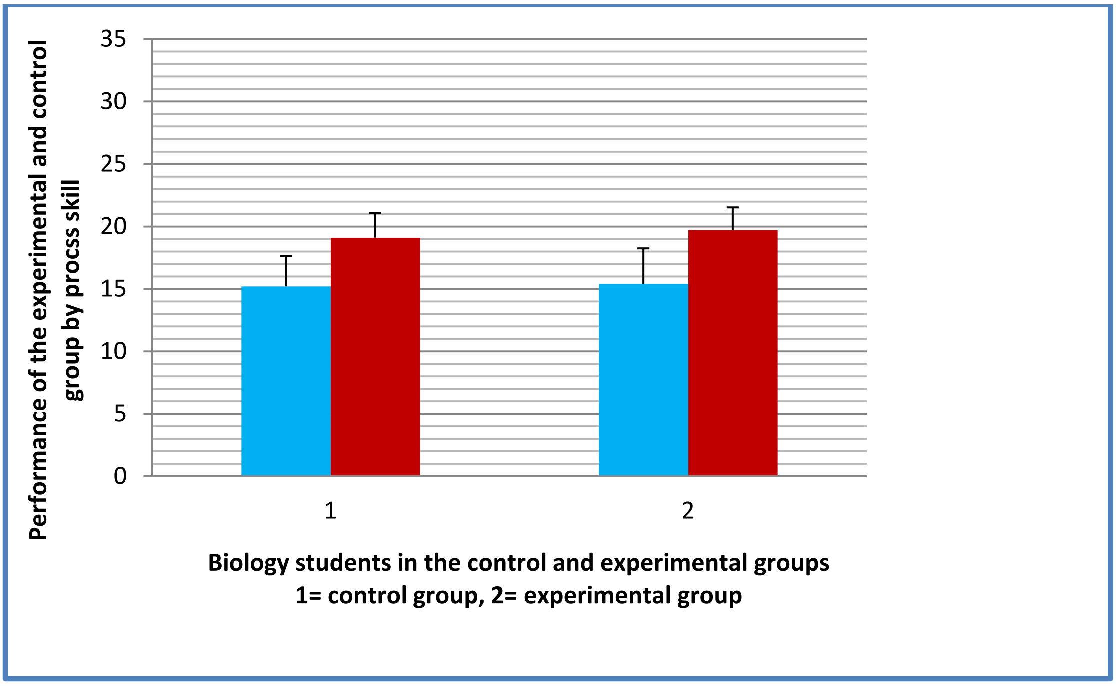
		Levene's test for equality of variances		*t-test for Equality of Means*						
									95% Confidence Interval of the Difference	
		F	*Sig.*	*t*	*df*	*Sig. (2-tailed)*	*Mean Difference*	*Std. Error Difference*	*Lower*	*Upper*
Students level of process skills	*Equal variances assumed*	1.620	0.204	0.633	261	0.047	0.15303	0.24158	-0.32266	0.62872
	Equal variances not assumed			0.619	179.795	0.037	0.15303	0.24712	-0.33460	0.64066

Source: *Field data (2015).*

As indicated in tables 6.5 (b), a t-test revealed a statistically significant difference between the mean score of control group students who were taught traditionally (M = 19.7, s.d = 1.822) and that of the experimental group(M = 19.5, s.d = 1.973), found a t (261) = 0.633, p = 0.047, α = 0.05 ($p<0.05$). It may be argued that students exposed to the inquiry based approach (the experimental group) had the opportunity to observe, discuss, interact and interpret data as they were. Hence it can be suggested that emphasis on students' participation in inquiry based lessons might have assisted the experimental groups to perform better in science process skills than the control groups students. It is in the view of this study that the teacher-centered mode of teaching science in the sampled schools, which did not allow the Biology students to practice and internalize the skills over a fairly long period. This might be one of the main reasons for the even experimental group students' poor performance on the many science process skills investigated.

Graph 6.1 below shows how the control and experimental group students performed science process skills test at pretest and at posttest occasions.

Graph 6.1: Performance of control group vs experimental group in the pretest and posttest (control group n= 94 & experimental group n= 169)



These results, however, support claims put forward by the Tanzania competence based syllabus (URT, 2005) that the inquiry-based method of teaching is more effective than the traditional method of teaching in science process skills development ad hence should be adopted by science teachers. On the other hand the findings from this study that experimental group students who were taught genetics using inquiry-based approach (IBA) outperformed control group students in science process skills resembles findings from many previous studies. For example, Lee and Butler (2003) examined the effect of designing and using inquiry tasks in increasing scientific knowledge and problem-solving skills. A sample of the study consisted of 59 male and female students who performed a set of real inquiry tasks (prediction, measurement, decision making). Results of the study indicated that the used teaching method was effective in promoting students' scientific understanding, enriching their knowledge base and their problem-solving ability which in turn contributes in preparing students to be active participants in the community.

The findings from the current study also resemble the findings put forward by Ghabayen (1982) who conducted a study to identify the effect on inquiry teaching method on preparatory school students' acquisition of physics concepts and scientific methods. A sample of the study consisted of (16) seventh-grade sections containing (228) male students and (340) female students assigned randomly into two groups: the first group was the experimental study group taught using the inquiry teaching method and the second group was the control and was taught using the traditional teaching method. The researcher used an achievement test and scientific methods test. Results of the study indicated that students in the experimental group students outperformed control group students in the physics concepts achievement test and in the acquisition of scientific methods.

Brian et al. (1994) conducted a study on a group of basic stage student teachers. A sample of the study was divided into (4) groups taught using (4) different teaching methods to identify the effect of each of these teaching strategies on students teachers acquisition for integrated science processes. Results of the study indicated that the cooperative learning group and lab activities based teaching method significantly outperformed students taught using the traditional teaching methods in acquiring scientific inquiry processes. However, this is not always the case that students exposed to inquiry-based teaching will always have good achievement in science process skills. Sometimes the acquisition of process skills is quite negative with an inquiry-based approach. For example, a study by German et al. (1996) examined and evaluated 7th-grade students' perceptions towards scientific inquiry processes skills. The study focused mainly on data recording, data analysis, data representation, findings representation and providing scientific evidence skills. A sample of the study consisted of (364) 7th-grade students and the Alternative Assessment of Science Process Skills (AASPS) to identify students' acquisition of scientific inquiry processes. Results of the study indicated that only (61%) of students were successfully able to perform the data recording related activities and that (69%) of students have not reached the required level in findings data representation skills in the designated activities. About 81% of students were not able to provide supportive scientific evidence to support the findings obtained in certain activities. However an exploration of the effect of directed inquiry approach integrated several learning strategies such as advance organizers, the learning cycle, concept maps, etc., on learning of science process skills by Germann (1989) reported that the directed inquiry approach to learning had no significant effect on the learning of science process skills or on cognitive development.

6.3.2.4 Compairing the performance of the control and experimental groups in the five BPST subscales understudy

Another objective of this study was to compare the performance of students in the experimental and control groups by specific process skills after the intervention. As it has already been stated in the earlier sections, the current study focused specifically on five scientific process skills namely i. formulation of hypotheses, ii. identifying and controlling variables, iii. design experiments iv. analyzing and interpreting data and v. defining variable operationally. Hence, the study was also interested at comparing the effectiveness of the inquiry-based approach and conventional teaching in the development of individual process skills to Morogoro students by comparing posttest BPST scores of the treatment group to those of the control group. Therefore, for additional analyses, the entire Biology process skills test posttest scores for the five subscales were descriptively and statistically analyzed. It has to be noted that each of the five subscales under study (process skills) was measured by 07 items in the BPST making a total of 35 questions.

The mean scores and standard deviations of the experimental group and control group on each BPST subscale were calculated using SPSS and summarized in table 6.6 and figure 6.2. The table indicates that correct response percentages for both groups (control and experimental) were highest for the process skills of hypothesis formulation with the mean score of 4.18 for the control group 4.41 for the experimental group. These posttest means of students are out of 07 total items for the subscale. On the other hand, the posttest means for both groups were lowest in the process skills of data analysis and interpretation. In this skill, the mean score of the control group was only 2.75 out of 07 while that of the experimental group was 2.37. The finding that both the control and experimental group did poorly in data analysis are in line with the findings by Foulds and Rowe (1996) who found that students were capable of identifying all variables influencing an experiment, scoring about 50% on the test items and they could also produce testable hypotheses, with scores of about 40%. However, they were unable to design a controlled experiment and analyze experiment results, gaining an average score of only 18% (Foulds and Rowe, 1996)

Apart from doing poor in questions measuring their ability in data analysis, descriptive statistics further indicates that students in both groups, students did poorly on the items measuring their ability to design experiments. The posttest mean for control group students was 3.54 and 3.13 for the experimental group students and control group students respectively. Furthermore, with defining operationally subscale as it is shown in Table 6.6 and figure 6.2 Morogoro Biology students in the control group had an average of correct responses of 3.74 while their counterparts in the experimental group had 3.67. The mean and standard deviations for the control group in controlling variable was 3.02 and 2.6 respectively while that of the experimental group was 3.00 and 2.5 in this subscale respectively as summarized in table 6.6 below.

Table 6.6(a) Posttest scores on BPST subscales based on the type of instruction students received (n=263)

Group Statistics					
	Type of Instruction	N	Mean	Std. Deviation	Std. Error Mean
Hypothesis formulation	Conventional approach	94	4.18	1.19	0.12
	Inquiry Based Method	169	4.41	1.31	0.10
Controlling variables	Conventional approach	94	3.02	2.61	0.26
	Inquiry Based Method	169	3.00	2.51	0.19
Designing experiments	Conventional approach	94	3.13	0.92	0.09
	Inquiry Based Method	169	3.54	0.84	0.06
Data analysis	Conventional approach	94	2.75	0.82	0.08
	Inquiry Based Method	169	2.37	0.66	0.05
Defining operationally	Conventional approach	94	3.74	0.97	0.10
	Inquiry Based Method	169	3.67	0.94	0.07

Source: Research survey (2015)

In an attempt to establish whether the mean scores of the control and experimental group observed in table 6.6 above were statistically significantly different, an SPSS independent sample t-test analysis was carried out and the results are shown on Table 6.6(b). Independent samples t-test was carried out for each of the five subscales posttest scores to determine the effectiveness of the inquiry approach of teaching vs traditional approach in each process skill development. According to the table 6.6(b), statistically significant difference was found in the hypothesis formulation subscale between the control and experimental group students' posttest scores. An analysis of independent samples t-test based on the type of instruction students received for this subscale at (α) =0.05 produced a p of 0.04 and a t value of -0.202, hence rejecting the null hypothesis (Ho). It means that there is statistically significant difference in students' development of hypothesis formulation ability between students exposed to inquiry-based teaching (IBA) and those exposed to traditional method (TM) in favor of the experimental group. Independent sample t-test of posttest scores also found a statistically significant difference between experimental group and control group students in terms of their ability in interpreting and analyzing data. An independent samples t-test based on type of instruction students received for design experiments subscale at (α) =0.05 found t (1,261) = 0.264, p=0.045. Hence a statistically significant difference between groups with respect to designing experiments in favor of the experimental group was noted rejecting the null hypothesis. On the other hand, an independent samples t-test based on the type of instruction students received for data analysis subscale at (α) =0.05 found t (1,261) = -1.789, p < 0.046. Hence a statistically significant difference between groups with respect to data analysis performance in favor of the control group was noted rejecting the null hypothesis (Ho) at 0.05.

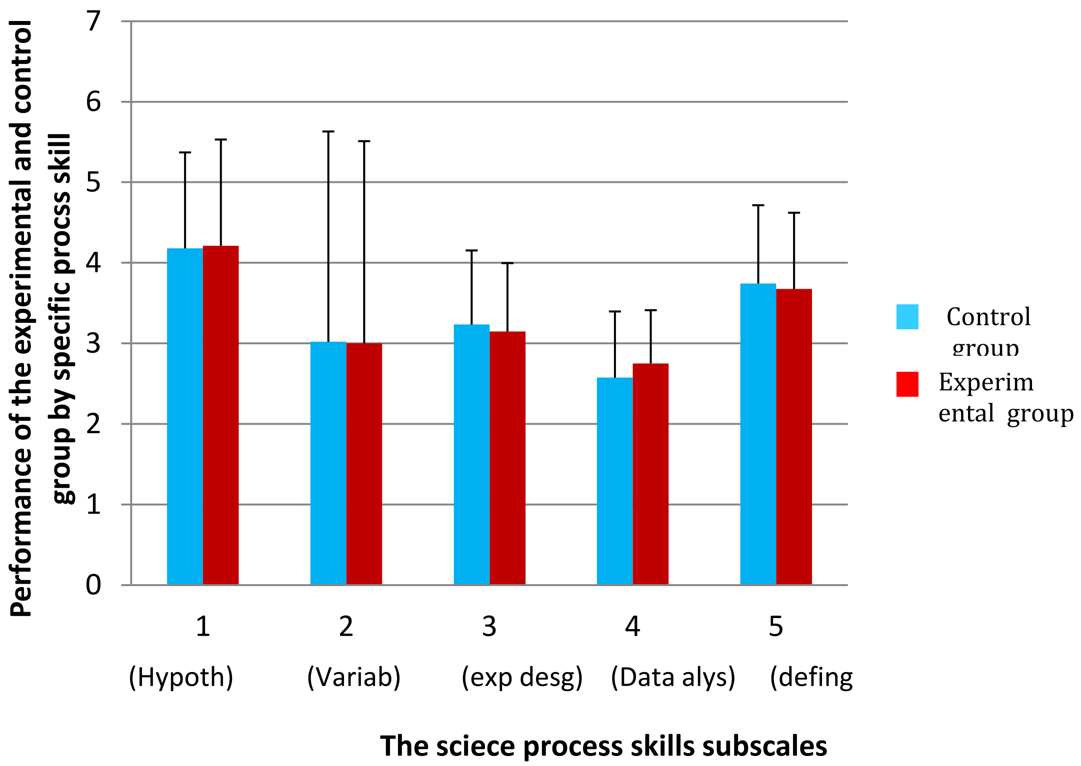
However, no statistically significant differences were found between students of control and experimental groups in the development of other science process skills understudy when their posttest scores were subjected to computer SPSS independent samples t-test. These skills included i. controlling variables and ii. defining operationally. This means that statistical analysis provides no evidence that those students who were taught genetics using inquiry method performed better in items measuring i. controlling variables, and ii. defining operationally than those who were taught genetics conventionally. These results, however, do not support claims put forward by the Tanzania competence based syllabus (URT, 2005) that the inquiry-based method of teaching is more effective than the traditional direct instruction method of teaching all science process skills development and hence should be adopted by science teachers.

Table 6.6(b): Independent samples t-test for posttest scores on the BPST subscales based of on the type of instruction students received (n=94 control group & n=169 experimental group)

		F	Sig.	t	df	Sig. (2-tailed)	Mean Differen	Std. Error Differenc	95% Confidence Interval of the Difference	
									Lower	Upper
Hypothesis form	Equ. variances assumed	5.660	0.018	-0.196	261	0.045	-0.032	0.164	-0.355	0.290
	Equ. variances not assumed			-0.202	209.26	0.040	-0.032	0.159	-0.346	0.281
Designing experm	Equ. variances assumed	1.125	0.290	0.264	261	0.045	0.086	0.112	-0.135	0.307
	Equ. variances not assumed			0.747	179.65	0.056	0.086	0.115	-0.141	0.313
Control Variables	Equ. variances assumed	.293	0.588	0.065	261	0.948	0.021	0.327	-0.624	0.666
	Equ. variances not assumed			0.064	186.10	0.949	0.021	0.331	-0.632	0.675
Data analysis	Equ. variances assumed	13.934	0.000	-1.902	261	0.038	-0.177	0.093	-0.360	0.006
	Equ. variances not assumed			-1.789	160.43	0.046	-0.177	0.098	-0.372	0.018
Operational defn	Equ. variances assumed	1.203	0.274	0.570	261	0.569	0.070	0.123	-0.172	0.312
	Equ. variances not assumed			0.566	188.32	0.572	0.070	0.123	-0.174	0.314

Source: Field data (2015)

Figure 6.2: Descriptive statistics of students score in the control group (n=94) and experimental groups(n=169) based on subscales of the BPST



Both the experimental and control group students in this study did poorly in items measuring their ability in data analysis and designing experiments (see table 6.6a above). Science education researches have highlighted a number of issues in the development of understanding in the areas of experimentation and science process skills. They argue that some of the scientific skills require a higher level of knowledge and skills use. One of these higher level skills is the skill of designing experiments and it may even be the skill that consists of all the other skills in its core. As seen in table 6.6, Biology students in Morogoro also performed poorly in items measuring their ability to design controlled experiments. According to Tan & Temiz (2003), this is the basic reason as to why students usually score poorly items measuring the skill of designing experiments. Experimenting skill is composed of building an appropriate mechanism by using various tools successfully; obtaining data by changing and controlling the variables, recording and assessing these data create models, interpreting the data, obtaining results and reporting the operations.

Another study by Bright & Friel (1998) found that when students first start working with graphs, they have difficulty moving back and forth between raw data for individuals and the group data represented in the graph. In his recent study Bülent (2015) examined preservice science teachers' skills of formulating hypotheses and identifying variables. At the end of his study, the results showed that preservice science teachers' skill at formulating a hypothesis and identifying dependent, independent and control variables accurately was low; their skill at identifying and controlling variables accurately was especially lower. Furthermore, the data from his observations indicated that pre-service science teachers had difficulty even in defining a hypothesis, formulating a hypothesis based on a problem, exemplifying ideal hypotheses, and defining, identifying and controlling variables. One of the most important reasons why preservice science teachers were not able to identify variables accurately was that they mistook or exchange variables. The cognitive and procedural skills associated with being able to select or conduct controlled experiments have been of interest to both science educators and psychologists who are interested in the development of scientific thinking (Schwichow et al. 2016). An investigative study by Schulz & Gopnik (2004) showed that students are able to select controlled experiments and to interpret unconfounded evidence when the experimental data are consistent with students' beliefs and preconceptions. Noss et al. (1999) studied use of data displays by practicing nurses and found that even though the nurses knew that blood pressure increases with age from their own experience and could use software to generate scatter plots of data on individuals' age and blood pressure, they were not able to "see" the relationship between the two variables in a scatter plot of these data.

6.3.2.5 Summary of findings on the effectiveness of inquiry-based approach on students' process skills development

This part discusses some key findings and their implications on the quasi experimental study to compare the effectiveness of inquiry-based teaching approach and conventional method on students' scientific process skills development. The analysis was carried out based on the hypotheses formulated, by using descriptive and inferential statistical tools revealed that the experimental group performed better after undergoing the experimental treatment of inquiry constructivist approach as compared to the scores of the control group. An analysis of independent samples t-test based on type of instruction students received at (α) =0.05 produced a p of 0.047 and a t value of 0.633, hence rejecting the null hypothesis (Ho1). The null hypothesis stated that, "there is no statistically significant difference in students' science process skills achievement between those exposed to inquiry-based teaching (IBA) and those exposed to traditional method (TM)". The null hypothesis was rejected at 0.05.

On the other hand, the general linear model with repeated measure for within group effects found $F(1,261) = 471.081$, $p < 0.001$, and eta squared =0.643. This means a significant within group effects was noted for time. This implies that regardless of the method of teaching in this study, there were significant within-groups effects with regard to the development of science process skills. The null hypothesis which stated that there is no statistically significant within group effects between control group and experimental group over two testing occasions (pretest and post test) with regard to students' science process skills development was rejected. Statistical significant time effects was noted at alpha =0.05 level. Lastly the test of between - subject effects found $F(1, 261) = 0.471.081$, $p < 0.157$ which means that the linear model accepted the null hypothesis. This further implies that the between- group interaction effects (method * groups* time) was not significant. Hence the null hypothesis was accepted at alpha = 0.05 level. This means that there were significant gains of science process skills to Morogoro students overtime but there were no statistically significant differential improvement among groups overtime. Table

6.6c is a summary of students' performance in the science process skills test based on the way they were taught changed from pretest to posttest.

Table 6.6.3 A summary of students' performance in the science process skills test (BPST) in the quasi experimental study (experimental group, n= 169; control group, n=94)

Table 6.6c A summary of students' performance in the science process skills test (BPST) in the quasi experimental study (experimental group, n= 169; control group, n=94)

		PRETEST RESULT VALUES			POSTTEST RESULT VALUES				
		Descriptive		**t-test**	**Repeated Measures**		**Descriptive**		**t-test**
		Mean	**s.d**		**Within F and η^2 values**	**Betwen F and η^2 values**	**Mean**	**s.d**	
Science process skills	contr	15.2	2.84	t=0.224	F=0.000 η2 = 0.643	F=0.157 η2 = 0.008	19.1	1.973	t=0.047
	exper	15.4	2.44				19.7	1.822	

Source: Research data (2015)

Independent samples t-test confirmed that inquiry based approach in this study was more effective than conventional approaches in acquiring science process skills. The significant difference in experimental group students' performances could be attributed to various direct experiences that gave participants the opportunity to question and formulate problems, manipulate materials, observe and record data, and reflect on and construct knowledge from the data. By reflecting scientific inquiry processes, the approach allowed students to become active participants in the process as they constructed an understanding of scientific concepts. These finding correlate the findings from a study which conducted by Roth and Roychoudhuri (1993) to examine the development of integrated science process skills in the context of open-inquiry laboratory as compared to normal traditional sessions. Findings from the study indicate that students develop higher-order process skills through nontraditional laboratory experiences that provided the students with freedom to perform experiments of personal relevance in authentic contexts. Students learned to, (a) identify and define pertinent variables (b) interpret, transform and analyze data (c) plan and design an experiment and (d) formulate hypotheses. The study suggests that process skills need not be taught separately. Integrated process skills develop gradually and reach a high level of sophistication when experiments are performed in meaningful contexts. Barman et al. (1996) attributed the success of the inquiry approach to providing opportunities for student interaction and dialogue through systematic instruction, learning experiences, and activities in each of the well-known phases. Because of the potency of the approach, students see the links among concepts explicitly and connect newly learned concepts to ones they already possess. To achieve meaningful acquisition of science process skills, learners must actively relate the ideas and facts that make up the concept. The strategies used in inquiry based classes supported a change in students from passively receiving information to actively examining their own conception.

Students in the experimental group were involved in activities that helped them reorganize their prior knowledge and develop scientific skills. They were allowed to think about their prior knowledge and reflect on it. This procedure helped students to learn meaningfully by making connections among concepts and by developing reasoning skills. Thus, students can learn facts, concepts, principles, and laws of science by directly engaging in science processes that require the use of their thinking abilities. In so doing, students will not only become proficient in the use of their thinking abilities, but also will remember and make sense of the associated concepts (Roth & Roychoudhuri, 1993). Beisenherz & Dantonio (1996) contend that by practicing the methods of science, students will not only better understand the nature of science and how it works but also develop thinking skills that will increase their ability to solve problems. While in the conventional based lessons, the teacher connected ideas for the learners in the inquiry based classes, students made the connections among the concepts by themselves through explorations and discussion. The important part in implementing the inquiry based instruction was the intensive teacher-student and student-student interaction because it provided students more time to discuss their findings with both their teacher and their peers.

Students who were taught genetics through the inquiry-based teaching approach attained higher scores in the BPST than those taught through the conventional method. This implies that using hands-on learning activities had a positive effect on students' development of science process skills. Education authorities in Tanzania should encourage science teachers to use this approach and teacher training institutions to make it part of their teacher training curriculum content.

Teacher training colleges and universities offering education courses should be designed to produce teachers capable of planning, designing and implementing inquiry-based teaching modules, lessons, and approach. Teachers in schools should be given training in planning and implementing inquiry-based teaching approach through in-service courses and orientations. This may be an effective teaching approach in providing suitable learning conditions for students of diverse learning styles and academic abilities that is common in most classroom settings. Students learn science best when the teaching methodology enables them to get involved actively in class activities. They should participate actively in doing experiments, carrying out demonstrations, class discussion and other relevant learning experience.

In this study, students who were taught genetics through the inquiry-based teaching approach attained higher scores in the BPST than those taught through the conventional lecture method. The results revealed that using hands-on learning activities had a positive effect on students' development of science process skills. Based on these findings, it can be concluded teachers in schools should be given training in planning and implementing inquiry-based teaching approach through in-service courses and orientations. This may be an effective teaching approach in providing suitable learning conditions for students of diverse learning styles and academic abilities that is common in most classroom settings. Students learn science best when the teaching methodology enables them to get involved actively in class activities. They should participate actively in doing experiments, carrying out demonstrations, class discussion and other relevant learning experience.

6.3.3 RESULTS WITH RESPECT TO STUDENTS' CONCEPTUAL UNDERSTANDING OF GENETICS

6.3.3.1 Pretest results from the Genetics test

Genetics is concerned with genes, heredity, and variation in living organisms. It seeks to understand the process of trait inheritance from parents to offspring, including the molecular structure and function of genes, gene behavior in the context of a cell or organism (e.g. dominance and epigenetics), gene distribution, and variation and change in populations. Another aim of this quasi-experimental study was to compare the effectiveness of the inquiry-based approach and traditional method of teaching in the students' conceptual understanding of content. Genetics was chosen as a case study because the topic is essentially a problem-solving science and offers a fruitful area for studying student problem-solving performance. The genetics pre-test was administered to the experimental group and the control group in order to determine whether the two groups of students were similar in terms of their genetics knowledge level before teaching intervention. The test measured five (05) subtopics in genetics as listed in the Tanzania Biology syllabus of 2010 for the advanced level students. These subtopics included i. hereditary materials (DNA/RNA), ii. genetic coding and protein synthesis, iii. Mendelian and non-mendelian inheritance, iv. sex- linked inheritance and pedigree analysis, and v. gene and chromosomal mutation. Lack or absence of significant differences between students´ pretest performance of the two groups would infer that the cognate abilities of the groups were approximately the same prior to the intervention. At this point the study intended to determine whether or not there statistically significant difference in conceptual understanding of genetics between students which are expected to be exposed to an inquiry-based teaching (IBT) approach and those expected to receive the traditional method (TM).

The current study involved 94 (35.7%) control group students who were taught themes of genetics by using the conventional (direct instruction) method and 169 (64.3%) experimental group students who were taught using inquiry-based approach (IBA). Boone (1990) suggested that when conducting teaching methodological studies with teachers delivering the treatments, precautions need to be taken to ensure conformity to teaching the approaches under investigation. Hence to ensure conformity in teaching that would provide a realistic comparison, the same instructor taught all course subtopics in both the control and experimental group.

The mean scores and standard deviations of the two groups in pretest are shown in Table 6.7 (a). It is noted that the genetics test composed of 25 multiple choice questions and it was marked with one point per each question. Hence the maximum score a student could score was 25 out of 25. The results show that the mean of scores of the experimental group was 9.8 out of 25 genetics questions with the standard deviation of 2.88, while the mean of the control group 9.6 out of 25 items and the standard deviation of 2.77. The results of the mean scores on genetics test are also represented in a bar graph in Figure 6.3. Spread (standard deviation) of individual scores around their respective means was 2.88 for the experimental group and 2.77 for the control group. This means that before intervention variability the experimental group (2.88) was more than that of the control group (2.77) as shown by the coefficient of variation. This could imply that the experimental group was more homogenous than the control group before teaching intervention. Many students failed to connect genes to proteins and phenotypes, and as a consequence fail to recognize the importance of proteins in this process, thus in some cases students incorrectly assume that genes are particles that directly express traits in organisms. In both groups, student scores ranged from 06 to 14 out of 25 items present in the test. Figure 6.3 in section 6.3.3.3 also summarizes pretest mean and standard deviation of the control group and experimental group students.

Table. 6.7(a) Group statistics for genetics pretest scores of students based on the type of instruction (n=94 control group & n=169 experimental group)

	Grade level of the students	*N*	*Mean*	*Std. Deviation*
Grade level of the students	Inquiry Based Approach	169	9.8	2.88
	Conventional Method	94	9.6	2.77

Source: Field data (2015).

An SPSS two-tailed independent samples t-test was conducted to test whether or not the observed pretest mean scores of experimental (IBA) and control (TM) classes on the genetics test are statistically significant or not. Table 6.7(a) indicates as if pretest performance of the experimental group as higher than that of the control group. However, no statistically significant difference was found between control and experimental group pretest mean scores on genetics test when the null hypothesis was subjected to the independent t-test. The null hypothesis stated that there is no statistically significant difference in the pre-conceptual understanding of genetics contents between those students to be exposed to inquiry-based approach and those to be exposed to traditional method (TM). An analysis of independent samples t-test based on genetics pretest mean scores of the experimental and control groups at alpha (α) =0.05 produced a p of 0.396 and a t value of 0.722. This means t-test failed to reject the null hypothesis at alpha (α) =0.05. Hence, the null hypothesis, that there is no statistically significant difference in the pre conceptual understanding of genetics contents between the control and the experimental group students was accepted at 0.05 alpha levels. Tables 6.7 (b) summarize the independent samples pretest t-test of both the control and experimental groups.

Table 6.7(b): Independent samples t-test for genetics pretest test scores (n=94 control group & n=169 experimental group)

		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
Students pretest scores in genetics	Equal variances assumed	0.722	0.396	-0.591	261	0.555	-0.21377	0.36192	-0.92642	0.49888
	Equal variances not assumed			-0.584	185.958	0.560	-0.21377	0.36604	-0.93590	0.50835

Source: *Field data (2015).*

The aim of administering genetics test before the actual intervention was to determine whether the experimental group and the control group were similar in terms of their pre-conceptual knowledge level genetics. These results from table 6.7 (a and b) above suggest that the pre-conceptual knowledge level of the genetics of the control and experimental group students were comparable prior to the genetics course intervention. This means that the groups exhibited comparable characteristics in terms of genetics content knowledge before the actual genetics course. Lack or absence of significant differences between the pretest performances of the two groups infers that the cognate abilities of the groups were approximately the same prior to the intervention. It was then concluded that these groups of Morogoro Biology students were suitable for the intended comparative study.

6.3.3.2 General linear model pretest posttest results comparison for the control and experimental groups

This study was also interested at examining the within and between group differences with respect to the development of students' conceptual understanding of genetics overtime. Therefore a repeated measures analysis of variance (ANOVAs) were conducted on students´ genetics scores to compare groups' performance over the two testing occasions and test for the between and within-group effects overtime. The measurement of time consists of time elapsed over 08 weeks of each aspect of study with measurement at pretest (week1) and post test (week 08). The findings from SPSS general linear model for repeated measure (within and between groups) are presented in section **6.3.3.2.1** and **6.3.2.3.2** below.

6.3.3.2.1 ANOVA for within - group differences (Test of within - subject effects)

Within-person (or within-subject) effects represent the variability of a particular value for individuals in a sample. In this study, a repeated measures analysis of variance (ANOVAs) was conducted on genetics scores to compare for the within-group differences overtime. Test of within - subject effects are an excellent measure to detect within-group differences over time. The intention was to test whether there is a significant mean gain score of the experimental and control group in genetics conceptual knowledge. In this test, the within subject factor was time with two levels (pretest in week 01 and posttest week in 08) and the dependent variables is the genetics scores at the pretest and posttest levels. Table 6.8(a) summarizes the findings of SPSS general linear model with repeated measure for pretest and posttest within- group effects with respect to genetics knowledge. A summarized in table 6.8 (a), the SPSS computation of general linear model with repeated measure for within -group effects (Sphericity Assumed) found $F(1,261) = 4.328$, $p < 0.001$, eta squared =0.943. Hence a significant main effect was noted for the time, $F(1, 261) = 4.328$, $p < 0.001$, which means regardless of the method of teaching there were a significant within groups effect on the conceptual understanding of genetics themes.

Table 6.8 (a): ***Within-subjects effects on genetics for two time periods (control group n= 94 & experimental group n= 169)***

			Tests of Within-Subjects Effects				
Measure							
Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Test scores (Genetics)	Sphericity Assumed	17564.207	1	17564.207	4.328E3	0.000	0.943

Source: *Field data (2015).*

This means that repeated measures analysis of variance rejected the null hypothesis that there is no statistically significant within-group effect in the conceptual understanding of genetics after teaching intervention over two testing occasions as a result of the methods of teaching. Eta square value was acquired as 0.943. This result shows that the effect magnitude is large and that almost 94.3% of the change in the dependent variable (genetic scores) results from the application. Student achievement increased in both groups as indicated by higher post-test scores. The experimental group increased their achievement but this was not statistically significantly different from the experimental group. This means Morogoro students learned the genetics content being taught in the same way regardless of teaching method and they perceived that student engagement was affected by the teaching method used.

6.3.3.2.2 ANOVA for between - subjects differences (Test of within - subject effects) in the Genetics test

A within subjects ANOVA was performed on genetics conceptual test scores to compare groups' scores over the two testing occasions. This multivariate repeated measures ANOVA was conducted with the factor being the two groups (control n= 94 and experimental group n= 169) overtime (pretest week and posttest week 8) and the dependent variable being student scores in the genetics conceptual test. The aim was to test statistically null hypothesis which stated that there is no statistically significant difference between control group students and experimental group students in the attainment of genetics knowledge over time. The general linear model for between- group interaction effects (method * groups* time) found F (1, 261) =0.924, p = 0.337. This means that the interaction was not significant at alpha = 0.5 and that the linear model accepted the null hypothesis. This means that there were significant gains over time and but there was no statistically significant differential improvement among the groups over time.

Table 6.8 (b): ***Between-subjects effects on genetics for two time periods (control group n= 94 & experimental group n= 169)***

			Tests of Within-Subjects Effects				
Measure:MEASURE_1							
Test and Type of instrction	Sphericity Assumed	3.750	1	3.750	0.924	0.337	0.004
Error(test)	Sphericity Assumed	1059.242	261	4,058			

Source: Field data (2015).

The statistical analysis revealed no significant difference in the performance of the two groups in the test, while both groups showed significant improvement (p<0.01) from the pre-test to the post-test. The findings in table 6.8 (b) implies further that regardless of the teaching method, there was an improvement of students genetics conceptual knowledge both, in the control and experimental groups. This means that both teaching methods (inquiry-based approach and conventional method) used in this study created a significant difference in advanced level high-school students' genetics disposition scores. Eta square value was acquired as 0.943. This result shows that the effect magnitude is large and that almost 94.3% of the change in the dependent variables (genetic scores) results from the application of the methods of teaching. The main findings showed that both methods had an impact on the development of genetics trend to students. These results, however, do not support anecdotal claims that the inquiry-based method of teaching is more effective than the traditional lecture method in enhancing the conceptual understanding of scientific concepts.

6.3.3.3 Posttest findings with the genetics test (Comparing the control and experimental groups)

Another purpose of this quasi-experimental study was to determine if there was a statistically significant difference in genetics achievement between experimental group students and the control group students. The overall aim here was to compare the effectiveness of the inquiry-based approach and conventional direct method in enabling conceptual understanding of Biology contents, with genetics being the case study. Student achievement was determined by the score comparison on 25 items multiple choice pre/post-test. Two independent-samples t-test was conducted to follow up the significant interaction and assess differences among teaching method groups at each time period. The hypothesis stated that there is no statistically significant difference in genetics achievement between students exposed to the inquiry-based mode of teaching (IBA) and those exposed to a traditional method (TM). The two groups were firstly given the pretest followed by a genetics intervention of 08 weeks before completing the same genetics test at posttest. The testing effects and influence of teacher variables across all the groups were nullified and the post-tests of each of the experimental groups could be compared with that of the control groups to detect the effects of an intervention (see section 6.2.4).

With the conventional method, lecture notes and discussion questions were prepared in advance before the actual class session. Three different textbooks prescribed by the Tanzania Biology syllabus and proved adequate to provide the essential factual basis for the course and were used in the construction of student's notes and discussion questions. They included Biological Sciences (1997) by D.J. Taylor, Understanding Biology for Advanced Level (1999) by Glenn Toole and Susan Toole, and Advanced Biology Principles and Applications by D.J Mackean and C.J Clegg. (2000) Each subunit met a total of 240 min/week (either 80 min on Monday/Wednesday/Friday or 120 min on Tuesday/Thursday) plus a 50-min recitation each week for a total of 8 weeks. In inquiry classes, many hours were dedicated to building new activities/models, and other activities. Throughout the teaching by inquiry, Biology students were working in small groups where they were encouraged to explore problems, formulate hypotheses, designing micro experiments share their ideas with their classmates, discuss their observations and interpret findings of the experiments or hands-on activity carried out. The school biology book was not used at all and the role of the teacher was reduced to that of a coordinator and facilitator of the students' work. The students' main learning aid was a set of worksheets which was collected from different sources mainly websites (see table 6.2 in section 6.2.5.2) prepared specifically for the teaching of the genetics. The worksheets complete with short articles as a source of new information, tables, diagrams, pictures, exercises, and guidelines for small investigations, facilitated the application of the inquiry approach. Experimental group underwent a total of sixteen inquiry-based lessons, of which two lessons on average were accomplished per week in eight weeks as shown in table 6.2 in section 6.2.5.2.

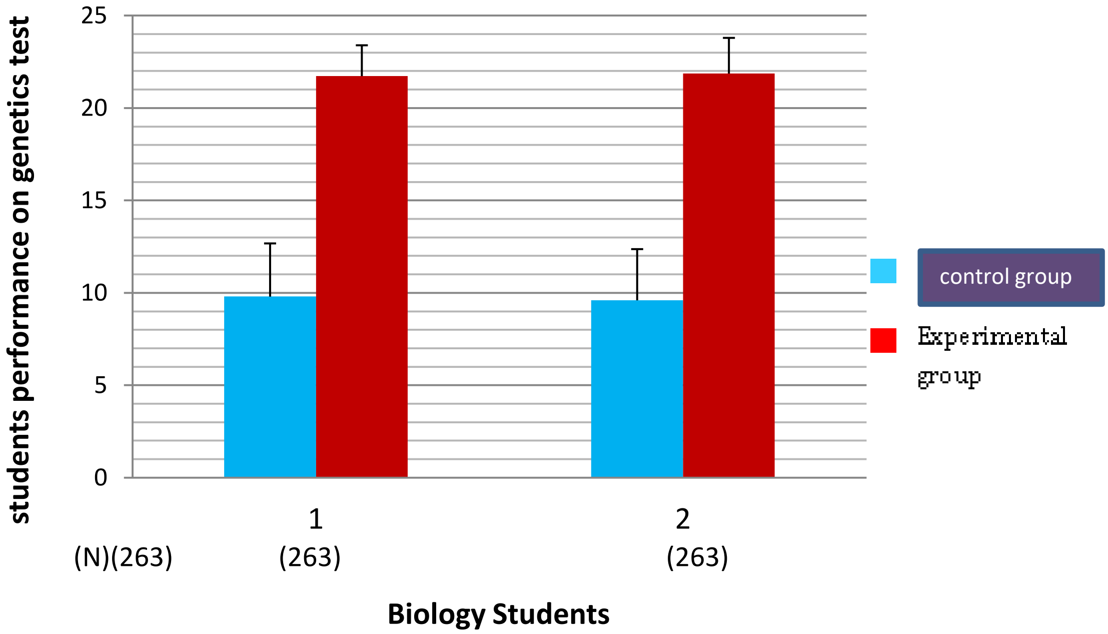
The posttest mean scores and standard deviations of the two groups on genetics test are shown in Table 6.9 (a). The results of the mean scores on genetics are also represented also in a bar graph in Figure 6.3. The mean of students score in the experimental group was 21.73 out of 25 questions with the standard deviation (sd) of 1.67, while the mean of the control group 21.87 out of 25 items with the standard deviation (sd) of 1.93. This means that from pretest in week one, the spread (standard deviation) of individual scores in the control group decreased from 2.77 to 1.93 and also decreased from 2.88 to 1.67 for the experimental group students. Contrary to pretest results, the variability the control group was more than that of the experimental group as shown by the coefficient of variation (1.93 for the control group and 1.67 for the experimental group). Hence the at the end of teaching intervention, the control group, in this case, was found to be more variable than the experimental group in terms of their genetics knowledge than their counterpart experimental group students. This means inquiry-based and hands-on activities that the experimental group students underwent made their genetics knowledge level more homogenous than the control group at posttest. The findings that students in the experimental were relative homogenous as compared to the control group are in line with the claim put forward by Keys & Bryan (2001) who argued that authentic inquiry activities provide learners despite their cognitive abilities with the motivation to acquire new knowledge, a perspective for incorporating new knowledge into their existing knowledge, and an opportunity to apply their knowledge.

Table 6.9(a) Group statistics for genetics posttest scores based on the type of instruction they received (n=263)

	Grade level of the students	*N*	*Mean*	*Std. Deviation*
Grade level of the students	Inquiry Based Approach	169	21.73	1.67
	Conventional Lecture Method	94	21.87	1.93

Source: Field data (2015).

Figure 6.3 Mean and standard deviations of the control and experimental group in genetics scores



A two tailed independent-samples t-test was conducted to statistically compare the posttest means of experimental (IBA) and control (TM) classes on the genetics test. The aim was to test whether or not the mean scores were statistically significant or not. As it has been indicated in table 6.9(a) above, at posttest the mean of scores of the control group was 21.87 out of 25 maximum while the mean of experimental group students was 21.73. However, no statistically significant difference was found between student posttest mean scores on the genetics test when the null hypothesis was subjected to independent samples t-test. The null hypothesis stated that there is no statistically significant difference in the posttest knowledge of genetics contents between students exposed to inquiry-based approach and those to be exposed to traditional method (TM). Independent samples t-test found the value for experimental group (M = 21.73. s.d = 1.67) and that of control group (M = 21.87, s.d = 1.93), t (261) = 0.606, p = 0.545, α = 0.05. Hence, the null hypothesis, that there is no statistically significant difference in genetics posttest scores between the control and the experimental groups was accepted at 0.05 alpha levels. Tables 6.9 (b) summarizes the independent samples pretest t-test of both the control and experimental groups.

Table 6.9(b): Independent samples t-test for genetics posttest scores (n=94 control group & n=169 experimental group)

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Students posttest scores in Genetics	Equal variances assumed	0.136	0.713	0.606	261	0.545	0.13861	0.22883	-0.311	0.589
	Equal variances not assumed			-0.580	169.007	0.563	0.13861	0.23903	-0.333	0.610

Source: *Field data (2015).*

As seen in table 6.9 (b), an analysis of independent samples t-test based on genetics posttest on experimental and control groups at alpha (α) =0.05 failed to reject the null hypothesis. This means that there were no statistically significant difference in the effectiveness of inquiry-based (IBA) approach and the conventional method (TM) in enhancing the conceptual understanding of genetics contents to students. These findings contradicts from the finding by Hadjimarcou et al. (2009) conducted a similar study to investigate the effectiveness of using an inquiry-based approach in teaching ninth-grade genetics in Cyprus. Their study involved teaching a unit of basic genetics to a control and an experimental group in the traditional teacher-centered and the inquiry approach, respectively. The results indicate that the inquiry method achieved a significantly better learning outcome compared to the traditional method. Leonard et al. (2001) found that students participating in a yearlong scientific inquiry-based Biology course posted higher gains in Biology concepts, and in the understanding of scientific processes. Furthermore, Alberts (2000) discovered that participating in scientific inquiry appears to improve retention of student learning. Leonard et al. (2001) found no differences in achievement in college chemistry between students who took an inquiry-based chemistry course in high school and those who took a traditionally taught chemistry course. The current findings do not resemble findings by Pinar & Ceren (2008) who also investigated the comparative effect of the learning cycle and expository instruction on 8th-grade students' achievement in genetics. The authors adopted the nonequivalent control group design as a type of quasi experimental design. The experimental group (n = 104) received learning cycle instruction, and the control group (n = 109) received expository instruction (conventional method). The learning cycle is an inquiry-based teaching strategy that divides the instruction into three phases: exploration, concept introduction, and concept application (Renner et al. 1988). The 2-way analysis of covariance indicated a statistically significant post-treatment difference between the experimental and control groups in favor of the experimental group after instruction.

However, as in similar studies (such as by Marbach, 2001; Lewis et al. 2000 and Pinar & Ceren, 2008), students' responses in the post-test items from the experimental group in this study revealed a number of difficult learning areas that students encounter in their effort to understand genetics. They include: i) the construction and interpretation of diagrams representing Mendelian inheritance, ii) the structure, function, and correlations between DNA, genes, and chromosomes, and iii) the way meiosis, mitosis, and fertilization collectively causes the appearance of the phenomenon of inheritance. Similar results also appear in other studies (Marbach, 2001; Lewis et al. 2000 and Pinar & Ceren, 2008). Watson et al. (1995) discovered that teachers used more extensive practical work in teaching science, while it had only a marginal effect on students' understanding of combustion.

6.3.4 FINDINGS ON MOTIVATION OF THE CONTROL AND EXPERIMENTAL GROUP TOWARDS SCIENCE PROCESS SKILLS

6.3.4.1 Introduction

This genetics teaching intervention was not designed specifically to contribute to students' attitudes towards science process skills. Yet, it was assumed that it might have an effect on the students' attitudes, perception and interest towards science. The competence based curriculum of Tanzania (URT, 2005) encourages teachers to use participatory inquiry and learning strategies as much as possible to help learners demonstrate self-esteem confidence and assertiveness. However, few studies if any have been conducted to assess the effectiveness of this approach (inquiry participatory) in the development of students' motivational levels as compared to traditional teaching. Hence another purpose of this quasi-experimental study was to compare the effectiveness of inquiry and traditional methods of teaching in developing students' motivation towards science process skills. In this respect, the students' motivation towards science process skills in both the experimental and control groups were scored prior to and after teaching intervention. The control group students and experimental group students were taught the topic of genetics differently in order to compare the level of students' motivational change towards science process skills at the end of teaching. Genetics was taught using the inquiry-based method to the experimental group students and by using the traditional method to the control group. The aim was to statistically answer the question, "what are the differences in students' motivation towards science process skills after they experienced inquiry and traditional science (genetics) instruction?

Science Motivation Questionnaire II by Glynn et al. (2011) and FSWEx Self-concept scale by Damerau (2012) were used as data collection tools. Science Motivation Questionnaire II by Glynn et, al. (2011) and FSWEx Self-concept scale by Damerau (2012) focuses on six motivational constructs which include i. intrinsic motivation, ii. grade motivation, iii. career motivation, iv. self-efficacy, v. self-determination, and vi. self-concept. The items on the Science Motivation Questionnaire II (SMQ) instrument had responses ranging from never to always. Its items were scored from 0, 1, 2, 3 and 4 for the responses and were scored in this manner: Always (A) = 4, Often (O) =3, Sometimes (S) =2, Rarely (R) =1 and Never (N) = 0. FSWEx self-concept questionnaire, on the other hand, consists of 18 items, which are further subdivided into three subscales, i. planning experiments (06 items), ii. practical experimentation, (06 items) and iii. analyzing data (06 items). The scale is based on the model of experimental skills (Schreiber et al. 2009) and uses a 5-point Likert scale ranging from 0 (strongly disagree), 1 (disagree), 2 (neutral), 3 (agree) to 4 (strongly agree). Contrarily from SMQ-II where the maximum score a student could score in a given subscale of five items was 20, in FSWEx self-concept questionnaire each subscale was measured with 6 items and the maximum score a student could get was 24 in each subscale.

In order to ensure conformity in teaching that would provide a realistic comparison, the same instructor taught all course sections included in the conventional method and inquiry-based approach groups. Students in both control and experimental groups take 20 minutes to fill out both the Science Motivation Questionnaire II by Glynn et al. (2011) and FSWEx Self-concept scale by Damerau (2012).

A mean score was calculated for each subscale of the - Science Motivation Questionnaire II pre-test and post-test in order to quantify the nature of the students' response to a specific motivation dimension which were intrinsic motivation and self-efficacy motivation. To establish whether the experimental and the control groups were similar in terms of their level of individual motivation towards science process skills at the beginning of the study, the pre-test scores of SMQ-II and FSWEx self-concept scale were descriptively analyzed and by using independent sample t-test. The total frequency and percentages of the students' response to individual questions items on the SMQ II and FSWEx self-concept scale were explored within the respective subscale, in order to identify any specific items that might have influenced the student to engage in science process skills. The following section represents pretest and posttest findings from these motivational Likert scales.

6.3.4.1 Comparing the groups in their intrinsic motivation (findings from SMQ-II subscale)

6.3.4.1.1 Pretest findings with regard to intrinsic motivation

Intrinsic motivation refers to the kind of motivation that comes from inside an individual rather than from any external or outside rewards, such as money or grades (Ryan & Deci, 2000). This study also intended to compare the effectiveness of inquiry approach to teaching and conventional method in developing intrinsic motivation of learners towards science process skills. Motivational researchers (Ryan & Deci, 2000; Grolnick & Ryan, 1987; Vallerand & Reid, 1984) have concluded that working on a task for intrinsic reasons is not only more enjoyable but also relates positively to learning, achievement, and perceptions of competence. To establish whether the experimental and the control group students had a similar level of intrinsic motivation towards science process skills at the beginning of the study, the pre-test scores of intrinsic motivation in the SMQ-II were analyzed descriptively and then by using independent sample t-test. According to Glynn et al. (2011) when using the Science Motivation Questionnaire II (SMQ II), students' raw scores on each scale should be converted to standard scores, establishing a derived scale consistent with the nature of the items. Standard scores according to Osterlind (2006) provide more practical information for decision making than raw scores. Hence standard scores Standard deviations mean and standard deviations were used for interpretation of scores.

The standard scores indicated that Morogoro students had low intrinsic motivation towards science process skills prior to intervention. The students in the treatment group for example obtained (M= 9.93, SD= 1.18) in the SMQ-II scale out of a possible 20 on this subscale compared to (M= 9.92, SD= 1.15) by the control group. Students' intrinsic motivation was assessed using the following statements

i. *Learning science process skills is interesting*
ii. *I am curious about discoveries in the processes of science*
iii. *The science processes I learn is relevant to my life*
iv. *Learning science makes my life more meaningful, and*
v. *I enjoy learning processes science*

The item statement number iii was scored high by students even before intervention (Mean, 2.41; SD, 0.493) while item statement one received the lowest score among the five items measuring student levels of intrinsic motivation. This means that although Morogoro students believed learning science processes to be relevant to life, they felt that learning them is not interesting. This item statement had a mean value of 1.76, and the standard deviation of 0. 476. More results are shown in table 6.10 below. The spread (standard deviation) of individual scores around their respective means was 1.15 for the control group and 1.18 for the experimental group. Standard deviation result indicates that experimental group was more heterogeneous with respect to the intrinsic motivation than the control group before teaching intervention.

Table. 6.10: Group statistics for intrinsic motivation pretest scores based on the type of instruction they received (n=263)

Group Statistics					
	Type of instruction	N	Mean	Std. Deviation	Std. Error Mean
Pretest intrinsic motivation scores	Conventional approach	94	9.92	1.15	0.119
	Inquiry Based Method	169	9.93	1.18	0.090

Source: Field data (2015).

However, no statistical significant differences in the level of intrinsic motivation towards science process skills were found between students of control and experimental groups when their pretest intrinsic motivation scores were subjected to computer SPSS independent samples t-test. Independent samples t-test for equality of means of intrinsic motivation pretest scores of the IBA (M=9.93, SD= 1.18) and that of TM classes (M=9.92, SD= 1.15); t (261) =0.570, p = 0.274, hence $p > 0.05$.The earlier hypothesis that the two groups do not significantly differ in terms of their science process skills was accepted. This implies that the two groups (control and experimental) had similar characteristics in respect to intrinsic motivation towards science process skills before the actual genetics course intervention and were therefore suitable for this comparison study.

6.3.4.2.2 Findings from the general linear model with repeated measures

The study also employed a general linear model with repeated measures in an attempt to examine the within-group effects and between-group effects with respect to the changes of students' intrinsic motivation over time. Therefore, a repeated measures analysis of variance (ANOVAs) was conducted on the scores from SMQ-II intrinsic motivation subscale over the two testing occasions to test for between and within-group differences over time. The measurement of the time consisted of time elapsed over 08 weeks of each aspect of study with measurement at pretest (week1) and post-test (week 08). In the test of between-subjects, the factor was the two groups (control n= 94 and experimental group n= 169) overtime (pretest week and posttest week 8) and the dependent variable was student scores in the SMQ-II intrinsic motivation subscale. In the test of within-subjects, the within-subject factor was time with two levels (pretest in week 01 and posttest week in 08) and the dependent variables is the SMQ-II intrinsic motivation scores at the pretest and posttest levels. The findings from SPSS general linear model for repeated measure (within and between groups) are presented in section 6.3.4.2.2.1 and 6.3.4.3.2.2 below.

6.3.4.2.2.1 ANOVA for within and between- group differences in intrinsic motivation (Test of within - subject effects)

Repeated measures analyses of variance (ANOVAs) were conducted on SMQ-II intrinsic motivation scores to compare groups change over the two testing occasions. It entails testing for between and within-group differences over time with regard to intrinsic motivation towards science process skills. The study compared statistically the effectiveness of inquiry-based teaching (IBA) approach and traditional lecture methods (TM) in the development of students' intrinsic motivation over time. As it has already been stated in section 6.3.4.2.2 above, in the test of within-subjects, the within-subject factor was time with two levels (pretest and posttest) and the dependent variable was student scores in the SMQ-II intrinsic motivation pretest and posttest.

Table 6.11(a) summarizes the findings of SPSS general linear model with repeated measure for pretest and posttest within and between groups (experimental and control groups). For time, SPSS computation of general linear model with repeated measure within groups (Sphericity Assumed) found F (1,261) = 4.86, p < 0.001, eta squared =0.943. Hence a significant main effect was noted for the time, F (1, 261) = 4.328, p < 0.001, which means regardless of the method of teaching, there was a significant within groups effect on the intrinsic motivation of students towards science process skills.

Table 6.11 (a): Within -subjects effects ANOVA on intrinsic motivation scores for two time periods (control group n= 94 & experimental group n= 169)

Tests of Within and Within-Subjects Effects							
Measure							
Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
SMQ-II intrinsic motivation)	Sphericity Assumed	7993.172	1	7993.172	4.860E3	0.000	0.949

Source: Field data (2015).

The null hypothesis that there is no statistically significant difference in the intrinsic motivation of students towards science process skills after teaching intervention within groups (control and experimental) over two testing occasions (pretest and post test) was rejected. Eta square value within groups was acquired as 0.949. This result shows that the effect magnitude is large and that almost 95% of the change independent variable (perceived intrinsic motivation) results from the application of methods of teaching.

6.3.4.2.2.2 ANOVA for between- group differences in intrinsic motivation (Test of within - subject effects)

The test for a between-subjects main effect of the teaching method of intrinsic motivation towards science process skills was conducted. Between-persons (or between-subjects) effects, examine differences between individuals or groups. However, the between-group interaction effects (method * groups* time) were not significant, F (1, 261) =0.241, p = 0.624. This means that the linear model with repeated measures accepted the null hypothesis for between-group effects. The hypothesis stated that with time there is no statistical significant between students exposed to inquiry-based teaching (IBA) approach and traditional method (TM) with respect to the level of their intrinsic motivation towards science process skills. This means that there were significant gains over time, but there was no statistically significant differential improvement among the groups over time.

Table 6.11 (b): Between -subjects ANOVA on intrinsic motivation scores for two time periods (control group n= 94 & experimental group n= 169)

Tests of Within and Within-Subjects Effects							
Measure							
test scores* type of instruction	Sphericity Assumed	,396	1	,396	,241	,624	,001

Source: Field data (2015).

The findings in table 6.11 (b) imply that regardless of the teaching method, there was an improvement of students intrinsic motivation towards science process skills both in the control and experimental groups. Different instructional practices are known to have different outcomes in student intrinsic motivation. For example, Ryan & Grolnick (1986) found that the more students perceived autonomy support in the classroom, the higher they reported self-worth, cognitive competence, internal control, and mastery motivation.

6.3.4.3 Posttest findings on intrinsic motivation (descriptive statistics and independent samples t-test)

Independent samples t-test was conducted to follow up the significant interaction and assess differences among teaching method groups with respect to their intrinsic motivation towards science process skills. The aim was to compare the effectiveness of inquiry-based approach vs traditional method on students' intrinsic motivation towards science process skills. The competence based curriculum of Tanzania emphasized inquiry-based approach should be an integral part of science teaching so that students acquire scientific skills and motivation. An analysis of SMQ-II intrinsic motivational subscale posttest mean scores was carried out. The null hypothesis stated that there is no statistically significant difference in students' intrinsic motivation towards process skills between students exposed to inquiry-based teaching (IBA) approach and those exposed to the conventional method (TM).

The mean of student scores and standard deviations of the two groups are shown in Table 6.12 (a). The data mean that the spread (standard deviation) of individual scores around their respective means changed from 1.15 to 1.45 for the control group and from 1.18 to 1.42 for the experimental group. As it was during pretest, the variability the control group (1.45) was more than that of the experimental group (1.42), even after posttest as shown by the coefficient of variation. The item statement i. "learning science process skills are interesting" still received lowest posttest mean scores of 3.15, while item statement iii. "the science processes I learn is relevant to my life" still received highest posttest scores with the mean of 3.53. Other item statements "I am curious about discoveries in the processes of science" received a posttest mean score of 3.18, "learning science process skills makes my life more meaningful" had a mean score of 3.28 and the last item "I enjoy learning processes science" had a mean score of 3.22.

Table 6.12 (a) SMQ-II (Intrinsic motivation) posttest scores based on the type of instruction (control group n= 94 & experimental group n= 169)

Group Statistics					
	Type of Instruction	N	Mean	Std. Deviation	Std. Error Mean
Posttest scores intrinsic motivation (SMQ-II)	Conventional approach	94	18.11	1.45	0.150
	Inquiry Based Method	169	18.01	1.42	0.104

Source: Field data (2015).

However, no statistically significant difference was found on students levels of intrinsic motivation towards science process skills based on the type of instruction they received during teaching when their SMQ-II (intrinsic motivation) posttest scores were subjected to SPSS independent samples t-test. The null hypothesis stated that "there is no statistically significant difference in students' intrinsic motivation towards process skills between students exposed to inquiry-based teaching (IBA) approach and those exposed to traditional method (TM)". An analysis of independent samples t-test based on the type of instruction students received (α) =0.05 produced a p of 0.569 and a t value of 0.570, hence accepting the null hypothesis at α =0.05 (table 6.12 (b).

Table 6.12(b): Independent samples t-test for SMQ-II Intrinsic motivation subscale posttest scores (control group n= 94 & experimental group n= 169)

		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
Students posttest scores in SMQ intr motv	Equal variances assumed	1.203	0.274	0.570	261	0.569	0.0121	0.036	-0.051	0.0937
	Equal variances not assumed			0.566	188.325	0.572	0.0121	0.037	-0.052	0.0944

Source: *Field data (2015).*

The findings that there is no statistically significant difference on students' level intrinsic motivation between those exposed to inquiry-based approach and those taught traditionally towards science process skills contradict a number of previous studies. However as Zoller (1991) found, not all students like new activity-oriented teaching. In fact, students' perceptions of cognitive demands and their appreciation of the new teaching model reflect the level of dissonance between their cognitive and affective styles with the new teaching model. In another study, Grolnick & Ryan (1987) found that non-controlling instruction resulted in greater interest and conceptual learning in students when compared with controlling instruction. Guthrie et al. (2000) describe an intervention that attempted to enhance the intrinsic motivation for the reading of students in the third and fifth grade. Instruction included autonomy support through self-directed learning, competence support in the form of strategy instruction, relatedness support in the form of student collaboration, learning goals, and the use of hands-on science activities like observation and data collection. Teachers emphasized learning goals and provided evaluative feedback on student work, but performance was not emphasized as a goal of learning. Students exposed to this instruction scored significantly higher with respect to curiosity and strategy use than students receiving traditional reading instruction but did not significantly differ in terms of extrinsic motivation indicators.

6.3.4.2 Comparing the groups in their perceived self-efficacy towards science process skills (findings from SMQ-II subscale)

Self-efficacy refers to an individual's belief in his or her ability to successfully perform a specific behavior. Students with a strong sense of self-efficacy approach difficult tasks as challenges to be mastered rather than as threats to be avoided. Fundamentally self-efficacy plays a central role in the extent to which individuals perceive their ability to master and feel competent about their ability to engage in specific behaviors (Bandura, 1997). Research shows that the type of learning environment and teaching method can improve self-efficacy in the classroom (Bandura, 1997; Fencl & Scheel 2005). This study also intended to compare the effectiveness of inquiry approach to teaching and conventional method in developing students' sense of self-efficacy towards science process skills. It was conceived that experimental group students taught genetics in the spirit of constructivism would develop a strong sense of efficacy towards processes of science than students in the control group who were taught traditionally. This section summarizes quasi-experimental findings from SMQ-II self-efficacy subscale. Part 6.3.4.3.1 presents pretest findings for the control and experimental groups, part 6.3.4.3.2 presents findings from the general linear model (repeated measures ANOVA) and lastly part 6.3.4.3.3 compares posttest levels of self-efficacy of control and experimental group student.

6.3.4.3.1. Findings from pretest on self-efficacy subscale

To establish whether the experimental and the control group students had a similar level of self-efficacy towards science process skills at the beginning of the study, the pretest scores of students on the self-efficacy subscale of SMQ-II were analyzed descriptively and then by using independent sample t-test. Students' self-efficacy towards science process skills was measured by the following statement items on the SMQ-II Likert scale,

i. I believe I can earn a grade of ''A'' in science process skills test,

ii. I am confident I will do well on science process skills tests,

iii. I believe I can master science process skills and knowledge and skills,

iv. I am sure I can understand science process skills,

v. I am confident I will do well on science process skills labs and projects.

Descriptive analysis of pretest scores on SMQ II self-efficacy subscale indicated that Morogoro students had low self-efficacy towards science process skills prior to the intervention. For example, students in the treatment group obtained an average (M) of 10.94 and standard deviation (SD) of 3.91 in their pre-test score out of a possible 20 on the subscale while the control group obtained (M= 10.67, SD= 3.64) out of 20. On the other hand, self-efficacy subscale scores showed that statement number three which read "I believe I can master science process skills and knowledge" was rated high by both the control and experimental students with the mean of 2.45 and standard deviation of 0.494. Statement number one on the other hand which read, "I believe I can earn a grade of ''A'' in science process skills test", received the lowest rating among the 05 items measuring student levels of self-efficacy. This item had a mean value of 1.18, and the standard deviation of 0. 986. Other statements in the questionnaire (ii. I am confident I will do well on science process skills tests, iv. I am sure I can understand science process skills, and v. I am confident I will do well on science process skills labs and projects) got moderate ratings from students. Table 6.13 below summarizes the pretest group statistics for self-efficacy scores.

Table 6.13 Group statistics for self-efficacy pretest scores based on the type of instruction they received (n=263)

	Type of instruction	N	Mean	Std. Deviation	Std. Error Mean
Self-efficacy subscale	Conventional approach	94	10.67	3.64	0.375
	Inquiry Based Method	169	10.94	3.91	0.301

Source: Field data (2015).

Despite differences in their self-efficacy mean scores, however, no statistically significant differences towards science process skills were found among students of control and experimental groups in the independent samples t-test. The Levene's test for equality of variances of self-efficacy pretest scores of the IBA (M= 10.94, SD= 3.91) and that of TM classes (M= 10.67, SD= 3.64); $t(261) = -0.550$, $p = 0.583$, hence $p > 0.05$.The earlier hypothesis that the two groups do not significantly differ in terms of their self-efficacy towards science process skills was accepted. This implies that the two groups had similar characteristics in respect to their self-efficacy towards science process skills and were therefore suitable for study.

6.3.4.3.2.1 Findings from the General linear model repeated measures ANOVA

Repeated measures analyses of variance (ANOVAs) were conducted on SMQ-II for self-efficacy pretest-posttest scores to compare groups change over the two testing occasions by comparing statistically the effectiveness of inquiry-based teaching (IBA) approach and traditional lecture methods (TM) in the development of self-efficacy of students' science process skills over time. The findings from repeated measures ANOVA for within and between-group effects with respect to students' self-efficacy findings are summarized in section 6.3.4.3.1 and section 6.3.4.3.2.

6.3.4.3.1 Repeated measures ANOVA for within group differences with respect to the sense of self efficacy towards science process skills

In the test of within-subjects, the within-subject factor was time with two levels (pretest in week 01 and posttest week in 08) and the dependent variables is the SMQ-II self-efficacy scores at the pretest and posttest levels. For the effects with time, SPSS computation of general linear model with repeated measure within groups (Sphericity Assumed) found $F (1,261) = 549.810$, $p < 0.001$, eta squared =0.678. Hence a significant main effect of self-efficacy was noted for the time which means regardless of the method of teaching there was a significant within groups effect on students' self-efficacy towards science process skills. Self-efficacy findings for within-subjects effects repeated measures ANOVA for two time periods are presented in table 6.14 below.

Table 6.14 (a) Within-subject ANOVA on self-efficacy for two periods of time (control group n= 94 & experimental group n= 169)

Tests of Between and Within-Subjects Effects							
Measure							
Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squa
SMQ-II self -effficacy test	Sphericity Assumed	2191.216	1	2191.216	549.810	0.000	0.678

Source: Field data (2015).

The null hypothesis that there is no statistically significant within-group effects between the control group and experimental group students on their self-efficacy level towards science process skills after teaching was rejected at alpha

= 0.05. This means teaching intervention with genetics brought significant within groups (control and experimental) over two testing occasions (pretest and post test). Eta square value within groups was acquired as 0.678. This result shows that the time-effect magnitude is large and that almost 67.8% of the change in self-efficacy results from the application of the methods of teaching (inquiry and traditional).

6.3.4.3.2 Self-efficacy scores for between-subject effects repeated measures ANOVA for two time periods

Repeated measures analysis of variance (ANOVAs) was conducted on SMQ-II for self-efficacy subscale scores to investigate for the between- group interaction effects (method * groups* time). In the test of between-subjects, the factor was the two groups (control n= 94 and experimental group n= 169) overtime (pretest week and posttest week 8) and the dependent variable was student scores in the SMQ-II self-efficacy subscale. According to Bandura (1997), educators should nurture students' self-direction and sense of self-efficacy by providing them with opportunities to exercise at least some degree of control over their own learning (constructivist and inquiry-based learning). However, no significant difference with respect to students self-efficacy was found between-group (experimental and control group) interaction effects with $F(1, 261) = 1.07$, $p = 0.004$. This means that linear model accepted the null hypothesis. The hypothesis stated that with time and treatment integrations, there are no statistically significant between-group effects with respect to self-efficacy towards science process skills. Overall, this implies that there were significant gains over time and but there was no statistically significant differential improvement among the groups over time. According to Schunk (1991), in any learning situation, students enter with a sense of efficacy that is based on their aptitudes and past experiences in similar tasks. Table 6.14(e) summarizes the findings of SPSS general linear model with repeated measure for pretest and posttest for within and between groups (experimental and control groups) with respect to self-efficacy levels.

Table 6.14(b) Between-subjects effects ANOVA on self-efficacy for two time periods (control group n= 94 & experimental group n= 169)

Tests of Between and Within-Subjects Effects							
Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squar
Test * Type of Instruction	Sphericity Assumed	4,265	1	4,265	1,070	,302	,004

Source: Field data (2015).

Different instructional practices are known to have different outcomes in student self-efficacy. However, the findings in table 6.14 (b) implies that regardless of the teaching method, there was an improvement of students self-efficacy motivation towards science process skills both, in the control and experimental groups. For example, Fencl & Scheel (2005) found in their study that pedagogies such as collaborative learning and inquiry-based activities have a strong correlation with how well students learn physics and their overall self-efficacy towards the subject. The students' response indicated that a question and answer format, inquiry-based lab activities and conceptual (rather than quantitative) problems had a significant effect on creating a positive climate in the classroom (Fencl & Scheel, 2005)

6.3.4.3.3 Posttest results: Comparing posttest results from self-efficacy subscale

Students' self-efficacy influences what they do, how hard they try, and how long they persist (Schunk, 1991). It has been established that teachers and their styles of teaching play a crucial role in instilling positive self – perceptions of efficacy in their students through training them to make use of a variety of learning strategies such as goal – setting, strategy training, modeling and feedback (Hattie, 2012; Schunk, 1995). Descriptive analysis and two independent-samples t-test were conducted to follow up the significant interaction and assess differences among teaching method groups with respect to their self-efficacy and towards science process skills. The aim was to compare the effectiveness of inquiry-based teaching approach vs traditional method on students' self-efficacy towards science process skills. The null hypothesis stated that there is no statistically significant difference in students' self-efficacy towards process skills between students exposed to inquiry-based (IBA) approach and those exposed to the conventional method (TM).

Descriptive analysis of posttest scores on SMQ II self-efficacy subscale indicates massive changes in both groups with respect to these motivational constructs at the end of teaching intervention. Table 6.9 (f) is a summary of group statistics for SMQ-II (self-efficacy) posttest scores of students based on the type of instruction they received. After the intervention, the heterogeneity of control group students in their self-efficacy towards science process skills decreased from 3.64 to 1.45 as shown by the standard deviation. The self-efficacy of students in the experimental group also increased from the mean score (M) of 9.79 and standard deviation (SD) of 0.74 in their pretest to posttest mean of 15.01 and standard deviation of 1.42. The experimental group variability with respect to self-efficacy towards science process skills also decreased from 3.91 to 1.42 as indicated by their standard deviation (scores around mean).

Students' self-efficacy towards science process skills was each used by 05 items present in SMQ II. For example, the SMQ-II self-efficacy subscale shows that item number five which asks students to indicate their scale on the statement I am confident I will do well on science process skills labs and project received the highest score. It had a mean of 3.78 and standard deviation of 1.18. In the pretest, the item number three, I believe I can master science process skills and knowledge was scored high by students (Mean, 2.45; SD, 0.494). Unfortunately item one, I believe I can earn a grade of ''A'' in science process skills test, still received lowest score among the five items measuring student levels of self-efficacy even after intervention. This means that the majority of Morogoro Biology students do not believe strongly that they can earn a grade of ''A'' in science process skills test set for them.

Table 6.15(a) Group statistics for self-efficacy posttest scores based on the type of instruction they received (n=263)

Group Statistics						
	Type of instr	N	Mean	Std. Deviation	Std. Error Mean
Posttest scores (SMQ II self efficacy subscale)	Conventional approach	94	15.1	1.45	0.15
	Inquiry Based Method	169	15.0	1.42	0.10

Source: Field data (2015).

Computer SPSS independent samples t-test was used to find out whether the mean scores of control and experimental groups in the self-efficacy subscale was statistically significant or not. No statistically significant difference was found on students self-efficacy posttest scores based on the type of instruction they received when the null hypothesis was subjected to independent samples t-test. SPSS independent samples t-test self-efficacy posttest scores of the IBA (M=15.01, SD= 1.42) and that of TM classes (M=15.11, SD= 1.45); found t (261) =0.957, p = 0.34, at alpha = 0.05. Hence the hypothesis that there is no statistically significant difference in students' self-efficacy towards science process skills between students exposed to inquiry-based teaching (IBA) approach and those exposed to traditional method (TM) was accepted.

Table 6.15(b): Independent samples t-test for SMQ- II (self efficacy) posttest Scores (control group n= 94 & experimental group n= 169)

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Students' self-efficacy	Equal variances assumed	0.045	0.832	0.957	261	0.340	0.03137	0.0320	-0.03321	0.09596
	Equal variances not assumed			0.959	193.854	0.339	0.03137	0.0327	-0.03314	0.09589

Source: *Field data (2015).*

These findings contradict a number of previous studies. For example, a recent study by Longo (2011) conducted an experimental study to determine the effect of inquiry-based instruction and specified that students learning in laboratory environments where research and inquiry-based activities are done had better self-efficacy perceptions compared to the students learning with traditional approaches. Bandura (1997) also concludes that cooperative learning strategies have the dual outcome of improving both self-efficacy and academic achievement. According to the author (Bandura, 1997), cooperative constructivist learning structures, in which students work together and help one another, also tend to promote more positive self-evaluations of capability and higher academic attainments than do individualistic or competitive ones. However, Zoller (1991) found that not all students like new activity-oriented teaching. In fact, students' perceptions of cognitive demands and their appreciation of the new teaching model reflect the level of dissonance between their cognitive and affective styles with the new teaching model. In another study, Grolnick & Ryan (1987) found that non-controlling instruction resulted in greater interest and self-efficacy plays a crucial role in science education. Self-efficacy beliefs are effective on students' actions regarding how much effort they expend on an activity and how long they put perseverance into an action when they face difficulties.

6.3.4.4 Comparing the perceived self-concept of students towards science process skills

Self-concept is an individual's awareness of her/his own identity. It is the cognitive aspect of self and generally refers to the totality of a complex, organized and dynamic system of learned beliefs, attitudes, and opinions that each person holds to be true about his or her personal existence (Lawrence, 1996). Student academic self-concept and its relations with other factors have been the focus of education and have attracted much attention over the past two decades (Abu-Hilal & Bahri, 2000). Studies have clearly demonstrated how important teaching approaches are, in influence the development of students' self-concept (Schweinhart et al. 1986). The argument is that the process leading to an enhancement of or decrease in the learner's self-concept begins with the interaction between teachers and students. In this respect, another purpose of this quasi-experimental study was to compare the effectiveness of inquiry and traditional methods of teaching in developing students' self-concept towards science process skills. The competence based curriculum of Tanzania (URT, 2005) encourages teachers to use participatory teaching and learning strategies as much as possible to help learners demonstrate self-esteem confidence and assertiveness. However, few studies if any have been conducted to assess the effectiveness of this approach (inquiry participatory) way of teaching in the development of students' self-concept compared to the traditional way of teaching. This is the essence of the current study.

This genetics teaching intervention was not designed specifically to contribute to students' self-concept towards science process skills. Yet, it was assumed that it might have a positive effect on the student attitudes. The FSWEx self-concept scale by Damerau (2012) was used as a data collection tool (see appendix VI). This questionnaire was designed to enable researchers and science teachers to gain a better understanding of the self-concept of students in doing science and to examine in which ways it affects the interest of doing science. FSWEx self-concept questionnaire consists of 18 items, which are further subdivided into three subscales, i. planning experiments (06 items), ii. practical experimentation, (06 items) and iii. analyzing data (06 items). The scale is based on the model of experimental skills (Schreiber et al. 2009) and uses a 5-point Likert scale ranging from 0 (strongly disagree), 1 (disagree), 2 (neutral), 3 (agree) to 4 (strongly agree). Contrarily from the SMQ-II questionnaire where the maximum score a student could have in a given subscale of five items was 20, in FSWEx Self-concept questionnaire each subscale was measured with six items and the maximum score a student could get was 24. Students' self-concept towards science process skills in both the experimental and control groups were scored prior to and after the genetics teaching intervention. This section summarizes quasi-experimental findings from FSWEx self-concept questionnaire. It is further subdivided into three parts, where part 6.3.4.4.1 presents pretest findings for the control and experimental groups, part 6.3.4.4.2 findings from the general linear model (repeated measures ANOVA) and part 6.3.4.3.3 compares posttest levels of the self-concept of control and experimental group students.

6.3.4.4.1 Pretest findings on the on the level of self-concept of students

To establish whether the previously described experimental and control group students had similar levels of self-concept towards science process skills, pretest scores on FSWEx questionnaire were analyzed descriptively and then by using independent sample t-test. The students were asked to rate statements measuring their self-efficacy by ticking either strongly disagree, disagree, neutral, agree or strongly agree depending on their feelings and perception. Planning experiments self-concept was measured by the following statements.

i. In my daily life, it often happens that questions emerge which can be solved by experiments.

ii. It is easy for me to formulate theoretically based hypotheses.

iii. It is quite easy for me to develop an experiment to solve a given problem.

iv. I am good at choosing suitable laboratory equipment for experiments.

v. It is easy for me to develop an experimental instruction to solve a specific scientific research question.

vi. I find it easy to transfer an idea for an experiment into an experimental setting.

Self-concept of students towards the experimenting, on the other hand, was measured by the following statements on FSWEx questionnaire.

i. I don´t have a good hand for carrying out experiments.

ii. I am good at working with laboratory equipment.

iii. Writing down experimental observation is always hard for me.

iv. I have no problem with arranging experimental setups.

I am good at doing experiments.

vi. Handling lab equipment is very easy for me.

Lastly, the self-concept levels of students towards analyzing experimental data were measured by the following statements.

i. Analyzing experimental data is easy for me.

ii. I often have problems with interpreting experimental results.

iii. Interpreting experimental observation is easy for me.

iv. I do well in analyzing experimental results.

v. Detecting possible errors in an experiment that went wrong is easy for me.

vi. I can easily generate graphs based on experimental data.

Descriptive analysis of pretest scores on FSWEx self-concept subscales indicated that in both groups, students had lower self-concepts towards analyzing experimental data than their self-concept towards planning experiments and the actual experimenting. In this subscale (data analysis), the 169 experimental group students had the mean score of 8.57 out of 24 possible while the 94 control group students had the mean score of 8.77 before the intervention. Furthermore, Morogoro students had better scores on self-concept towards planning experiments than their self-concept towards actual experimentation and analyzing data subscales. The result means that Morogoro students had a positive picture of their ability to plan experiments than doing actual experiments and analyzing data before intervention (table 6.16 (a).

Table 6.16 (a) pretest scores on FSWEx self-concept subscales (Planning, experimenting & analyzing) (n=263)

Group Statistics					
	Type of Instrucion	N	Mean	Std. Deviation	Std. Error Mean
FSWEx(planning experiment subscale)	Conventional approach	94	10.65	2.24	0.23
	Inquiry Based Method	169	10.80	2.30	0.17
FSWEx (experimenting subscale)	Conventional approach	94	9.53	3.13	0.32
	Inquiry Based Method	169	9.20	2.96	0.22
FSWEx (Data analysis subscale	Conventional approach	94	8.77	3.15	0.32
	Inquiry Based Method	169	8.57	3.14	0.24

Source: Field data (2015).

In the planning experiment subscale, the item statement number iv. I find it easy to transfer an idea for an experiment into an experimental setting had lowest mean scores with 1.22 out of 4.0 possible. This means that before intervention students in both groups had the feeling that it is difficult for them to transfer an idea for an experiment into an experimental setting. The item number one in this subscale, (in my daily life it often happens that questions emerge, which can be solved by experiments) was higher rated by Morogoro students of both groups than any other items. It had the mean of 3.34 even before the intervention. On the other hand, scores in the data analysis subscale indicated that item number iv. (detecting possible errors in an experiment that went wrong is easy for me) had lowest scores with a mean of 1.96 and standard deviation of 0.411. This implies that before intervention students in both groups perceive themselves as incapable of detecting errors in experiments.

However, it was also necessary to find out whether the mean of the control and experimental groups observed above differed statistically or not. It has to be noted that these are pretest scores (scores before intervention). Hence the SPSS independent samples t-test was used with the FSWEx self-concept subscales pre-test scores for i. planning experiments, ii. practical experimentation and iii. data analysis. In both cases, however, no statistically significant differences between the experimental and control groups were found in students' sense of self-concept pretest scores when null hypothesis was subjected to independent samples t-test. For planning experiments subscale for example, Levene's test for equality of variances of IBA classes (M=10.65, SD= 2.24) and that of TM classes (M=10.8, SD= 2.30); found $t(261) = -0.494$, $p = 0.622$, at alpha = 0.05. Hence the null hypothesis that regardless of their groups, Morogoro students did not differ significantly in their sense of ability towards planning science experiments before the intervention was accepted. Levene's test for equality of variances for pretest scores of practical experimentation subscale also failed to reject the null hypothesis at $\alpha = 0.05$. Independent samples t-test for practical experimentation subscale pretest scores of the IBA (M=9.20, SD= 2.96) and that of TM classes (M=9.53, SD= 3.13); produced $t(261) = 0.834$, $p = 0.405$, at alpha = 0.05. Hence the null hypothesis that there is no statistically significant difference in the perceived ability for practical experimentation between the control and experimental group students was also accepted. Table 6.16(b) below summarizes findings from t-test for equality of means from pretest scores of both groups at α = 0.05 (all three FSWEx subscales).

Table 6.16 (b) Independent Samples t-test on self-efficacy pretest scores (Planning experiments, experimenting a& analyzing data) (n=263)

		F	Sig.	t	df	Sig. (2-tailed)	Mean Differe	Std. Error Differ	95% Confidence Interval of the Difference	
									Lower	Upper
preFSWE PLANING	Equal variances assumed	0.217	0.642	-0.494	261	0.622	-0.145	0.293	-0.723	0.433
	Equal variances not assumed			-0.498	197.08	0.619	-0.145	0.291	-0.719	0.429
preFSWE XPERI	Equal variances assumed	0.266	0.606	0.834	261	0.405	0.324	0.389	-0.441	1.09
	Equal variances not assumed			0.821	183.14	0.413	0.324	0.395	-0.456	1.10
preFSWE ANALYS	Equal variances assumed	0.062	0.804	0.501	261	0.617	0.202	0.404	-0.594	.999
	Equal variances not assumed			0.500	191.78	0.618	0.202	0.405	-0.596	1.00

Source: Field data (2015).

Lastly, as shown in table 6.16(b) above, Levene's test for equality of variances for pretest scores of analyzing data subscale also failed to reject the null hypothesis at α = 0.05. Independent samples t-test for data analysis subscale pretest scores of the IBA (M=8.57, SD= 3.14) and that of TM classes (M=8.77, SD= 3.15), produced t (261) =0.501, p = 0.617, at alpha = 0.05. Hence the null hypothesis that there is no statistically significant difference in the sense of ability to analyze and interpret experimental data between students exposed to inquiry-based teaching (IBA) approach and those exposed to traditional method (TM) before the intervention was also accepted. These findings imply that the two groups (experimental and control) of Morogoro students that were involved in this quasi-experimental study had a similar sense of ability both in planning experiments, actual experimenting and in data analysis. The two groups were, therefore, suitable for study.

6.3.4.4.2 General linear model for self concept pretest posttest results (comparison of the control and experimental groups)

Repeated measures analysis of variance (ANOVAs) was also conducted FSWEx self-concept subscales posttest scores to compare groups change in the sense of their ability towards planning experiments, actual implementation and data analysis over the two testing occasions. This measure analyzed both the between and within-group differences overtime with regard to students self concept changes towards science experiments (the science process skills). The study intended to compare statistically the effectiveness of inquiry-based approach and traditional method in the development students' self-perception of their ability in i. planning experiments, ii. actual implementation and in iii. data analysis (the processes of science) overtime. In the test of between-subjects the factor was the two groups (control n= 94 and experimental group n= 169) overtime (pretest week and posttest week 8) and the dependent variable was student scores in the FSWEx self-concept subscales. In the test of within-subjects, the within subject factor was time with two levels (pretest in week 01 and posttest week in 08) and the dependent variables was the FSWEx self-concept subscales scores at the pretest and posttest levels. The computer SPSS version 21 general linear model with repeated measure was used for comparing the pretest and posttest scores for within and between groups effects with respect to students change in their sense of self concept towards science experimentation. This part presents the findings from general linear model with repeated measures for the control and experimental group students. Part 6.3.4.4.2.1 presents findings from planning experiments subscale, part 6.3.4.4.2.2 are findings on students' self-concept towards practical/ actual experintation and lastly part 6.3.4.4.2.3 which present findings on students' self-concept towards data analysis and interpretation from scientific practicals**.**

6.3.4.4.2.1 Repeated measures analysis of variance (ANOVAs) for within and between group effects self-concept on planning experiments

For the effect of time, SPSS computation of general linear model with repeated measure within groups on planning experiments FSWEx subscale pretest and posttest scores (Sphericity Assumed) found $F(1,261) = 1285.156$, $p < 0.001$, eta squared =0.831. Hence a significant effect on students' changes in their self-concept towards planning experiments was noted with time. This means that regardless of the method of teaching used (inquiry and conventional) there was a significant within groups effect on students self-concept towards science experiments (science process skills). The null hypothesis which stated that with time, there were no statistically significant within-group differences between control group and experimental group students in their self-concept level towards science experimentation was rejected at alpha = 0.05. This means that the intervention brought significant within groups differences (control and experimental) over two testing occasions (pretest and post test). Eta square value within groups was acquired as 0.831. This result shows that the time-effect magnitude is large and that almost 83% of the change in self-concept results from the application of the methods of teaching (inquiry and traditional). Table 6.17 (a) below presents the findings from the test of within-subjects on scores from the self-concept on planning experiments subscale of FSWEx.

Table 6.17 (a) ANOVA for within and between-subjects effects on self-concept towards planning experiments *(control group n= 94 & experimental group n= 169)*

Measure							
Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
FSWEx scores	Sphericity Assumed	3352.393	1	3352.393	1285.156	0.000	0.831
FSWEx scores * type of instr	Sphericity Assumed	0.636	1	0.636	0.244	0.622	0.001

Source: Field data (2015).

Repeated measures analysis of variance (ANOVAs) was also conducted on FSWEx (planning experiment subscale) pretest and posttest scores to investigate and compare the between- group interaction effects (method * groups* time). The aim was to answer the question, is there the statistical significance between groups effects in their self-concept towards the ability of planning experiments after they had undergone different method of teachings? However, no significant difference in student' self-concept towards planning experiments was found between- group (experimental and control group) with F (1, 261) =0.224, p =

0622. This means that linear model accepted the null hypothesis at alpha = 0.05. The hypothesis stated that with time and treatment interactions, there was no statistically significant between-group difference with respect to their self-concept towards planning experiments (science process skills). Overall, this implies that there were significant gains over time but there was no statistically significant differential improvement among the groups over time. Hence, regardless of the teaching method employed, there was an improvement of students' self-concept towards their ability to plan science experiments in both, the control and experimental groups.

6.3.4.4.2.2 Repeated measures analysis of variance (ANOVAs) for within and between group effects self-concept on actual experimenting

In comparing the effectiveness of inquiry and traditional method in enhancing students self-concept towards undertaking practical experimentation over time of intervention, the FSWEx experimenting subscale posttest scores were analyzed. For the effects of treatment with time, the general linear model with repeated measure within groups (Sphericity Assumed) found $F(1,261) = 1669.541$, $p < 0.001$, eta squared =0.865. Hence a significant effect on students' changes in their self-concept towards actual experimentation with time was noted. This means that in both groups, the teaching method used (inquiry for the experimental group and conventional control group) had a significant within-group effect on students' self-concept towards experimenting (science process skills). The null hypothesis which stated that with time, there was no statistically significant within-group effects between the control group and experimental group students on their self-concept towards experimenting was rejected at alpha = 0.05. This means that the intervention brought significant within groups differences (control and experimental) over two testing occasions (pretest and posttest). Eta square value within groups was acquired as 0.865. This result shows that the time-effect magnitude is large and that almost 86.5% of the change in students self-concept towards their ability in experimenting results from the application of the methods of teaching (inquiry and traditional). Table 6.17 (b) summarizes findings from the repeated measures ANOVA for within and between subjects effects on the self-concept of both the control and group students towards conducting experiments.

Table 6.17 (b) ANOVA for within and between-subjects effects on self-concept towards conducting experiments *(control group n= 94 & experimental group n= 169)*

Tests of Within-Subjects Effects							
Measure							
Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
FSWEx scores	Sphericity Assumed	15113.962	1	15113.962	1669.541	0.000	0.865
FSWEx * type of instruction	Sphericity Assumed	5.696	1	5.696	0.629	0.428	0.002

Source: Field data (2015).

Repeated measures analysis of variance (ANOVAs) was also conducted on FSWEx (experimenting subscale) posttest scores to investigate and compare the between- group interaction effects (method * groups* time). The aim was to answer the question, is there statistically significant between groups differences in their self-concept towards the ability to experiment after they had undergone different teaching methods. No significant difference in student' self-concept towards experimenting was found between- group (experimental and control group) with F (1, 261) =0.629, p = 0428 at eta squared of 0.002. This means that linear model accepted the null hypothesis at alpha = 0.05. The hypothesis stated that with time and treatment interactions, there was no statistically significant between-group difference with respect to their self-concept towards experimenting (science process skills). Overall, this implies that there were significant gains in the self-concept towards experimenting over time, but there was no statistically significant differential improvement among the groups over time. Hence regardless of the teaching method employed, there was an improvement of students' self-concept towards their ability to experimenting in both the control and experimental groups.

6.3.4.4.2.3 Repeated measures analysis of variance (ANOVAs) for within and between group effects self-concept on data analysis

Data analysis and interpretation was one of the important kinds of science process skills focused in this study. It entails the ability of students in assigning meaning to the collected information and determining the conclusions, significance, and implications of the experimental findings. In comparing the effectiveness of inquiry vs and traditional methods in enhancing students' self-concept towards undertaking analyzing experimental data, the FSWEx data analysis subscale pretest and posttest scores were analyzed. For the effects of treatment with time, the general linear model with repeated measure within groups (Sphericity Assumed) found $F(1,261) = 3446.789$, $p < 0.001$, eta squared =0.93. Hence a significant effect on students' changes in their self-concept towards analyzing data with time was noted. This means that in both groups, the teaching method used (inquiry and control group) had a significant within-group effect on students' self-concept towards analyzing data (science process skills). The null hypothesis which stated that with time, there was no statistically significant within-group differences between control group and experimental group students on their self-concept towards analyzing data was rejected at alpha = 0.05.

This means that the intervention brought significant within groups differences (control and experimental) over two testing occasions (pretest and post test). Eta square value within groups was acquired as 0.93. This result shows that the time-effect magnitude is large and that almost 93% of the change in students self-concept towards their ability in analyzing data results from the application of the methods of teaching (inquiry and traditional). Table 6.17 (c) summarizes findings from the repeated measures ANOVA for within and between subjects effects on self-concept towards analyzing experimental data.

Table 6.17 (c) ANOVA for within and between-subjects effects on self-concept towards analyzing data (control group n= 94 & experimental group n= 169)

Tests of Within-Subjects Effects							
Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
FSWEx scores	Sphericity Assumed	18395.86	1	18395.8	3446.78	0.000	0.930
FSWEx * type of instruction	Sphericity Assumed	0.126	1	0.126	0.024	0.878	0.000

Source: Field data (2015).

As seen in table 6.17 (c) above repeated measures analysis of variance (ANOVAs) was also conducted on FSWEx (analyzing data subscale) posttest scores to investigate and compare the between- group interaction effects (method * groups* time). The aim was to answer the question, is there statistically significant between groups differences in their self-concept towards the ability to analyze data from experiments after they had undergone different teaching methods. No significant difference in student' self-concept towards analyzing data was found between- group (experimental and control group) with F (1, 261) =0.24, p = 0.878. This means that linear model accepted the null hypothesis at alpha = 0.05. The hypothesis stated that with time and treatment interactions, there was no statistically significant between-group difference with respect to their self-concept towards analyzing experimental data (science process skills). Hence, regardless of the teaching method employed, there was an improvement of students' self-concept towards their ability to analyze data in both, the control and experimental groups.

These findings that the control group and the experimental group did not differ significantly in their self-concept towards planning experiments, actual experimentation analyzing data resembles the conclusion put forward some researchers. For example Suk Kim (2005) conducted a study on the effects of a constructivist teaching approach on student academic achievement, self-concept and learning strategies and concluded that constructivist teaching is not effective in terms of student self-concept enhancement and student learning strategy changes in general, but have some effect upon motivation to learn academic tasks, causing anxiety in the academic learning process and self-monitoring in terms of learning for tests.

6.3.4.4.3 Findings on Posttest Scores: Comparison of the Control and Experimental group students on self-concept towards science experimentation (Science process skills)

Descriptive analysis and two independent samples t-test were conducted to follow up the significant interaction and assess differences among teaching method groups with respect to their self-concepts towards experimentation in science. The aim was to compare the effectiveness of inquiry-based teaching approach vs traditional method on students' self-concept towards science process skills. This part of the study compared the effectiveness of inquiry approach with the traditional method in developing the positive self-concept of students towards i. planning experiments ii. actual experimenting, and iii. analyzing experimental data. The null hypothesis stated that there is no statistically significant difference in students' self-concept between? those exposed to inquiry-based teaching (IBA) approach and those exposed to the conventional traditional method (TM).

Descriptive analysis of posttest scores on FSWEx self-concept subscales indicates a massive change in both groups with respect to the investigated concepts at the end of teaching intervention. Table 6.18(a) is a summary of group statistics for FSWEx questionnaire posttest scores for all three subscales of students based on the type of instruction they received. The planning experiment subscale posttest scores in both groups show that the item number i. in my daily life, it often happens that questions emerge which can be solved by experiments and item number ii. It is easy for me to formulate theoretically based hypotheses to have the highest mean scores than other items. After the intervention, these items had mean scores of 3.84 and 3.65 respectively. Contrarily to pretest, in posttest the item number v, it is easy for me to develop an experimental instruction to solve a specific scientific research question had the lowest mean score than other items. The item still had a mean score of 2.25 and standard deviation of 0.478. Table 6.18(a) summarizes the mean score and standard deviation for FSWEx planning experiment subscale posttest.

For changes in self-concept towards experimenting from pretest to posttest, the table indicates, for example, the 94 control students obtained (M= 21.14, SD=

1.49) in their post-test scores out of a possible 24 compares to (M= 9.53, SD= 3.13) in their pretest scores. The mean scores of 169 experimental group students also rose from (M= 9.2, SD= 2.96) in their pretest to (M= 20.88, SD= 1.43) after teaching genetics using inquiry-based approach. In the experimenting subscale, posttest scores in both groups show that the item number v. handling lab equipment is very easy for me and item number ii. i am good at working with laboratory equipment to have the highest mean scores than other items. After the intervention, these items had mean scores of 3.75 and 3.73 respectively. Contrary to pretest, in posttest the item number v, I am good at doing experiments had the lowest mean score than other items. The item still had a mean score of 2.43 and standard deviation of 0.485 even after posttest. Table 6.18(a) below have a summary of the mean score and standard deviation for FSWEx questionnaire experimenting subscale posttest results.

Changes in the self-concept of Morogoro Biology students towards their ability in analyzing experimental data were also assessed. The mean scores of 169 experimental group students also rose from (M= 8.57, SD= 3.14) in their pretest to (M= 18.28, SD= 1.27) after teaching genetics using inquiry approach. In data analysis subscale, posttest scores in both groups show that the item number iv. I do well in analyzing experimental results and item v. detecting possible errors in an experiment that went wrong is easy for me to have the lowest mean scores than other items. Even after the intervention, these items had mean scores of 2.12 and 2.13 respectively. Item number v, I am good at doing experiments had the lowest mean score than other items. Biology students, on the other hand, scored high the item number vi (I can easily generate graphs based on experimental data) with a mean score of 3.11 and standard deviation of 0.761 as seen in the table 6.18 (a) below.

On the other hand, the heterogeneity of control group students in their self-concept towards their ability to plan science experiments decreased from 2.24 to 1.55 and of the experimental group from 2.30 to 1.58 as shown by the standard deviation. This means that at the end of intervention students were more uniforms in terms of their self-concept towards planning experiment. This is also true to other self-concept variables assessed in this study.

For example, the standard deviation of the control group in FSWEx questionnaire (experimenting subscale) decreased from 3.13 in their pretest to 1.49 after intervention (posttest) while that of the experimental group also decreased from 2.96 to 1.43 after the intervention. In the data analysis subscale, posttest results show that the heterogeneity of the control group decreased from 3.13 to 1.26 while that of the control group also decreased from 3.14 to 1.27 after the intervention.

Table 6.18 (a) Students FSWEx posttest scores (control group, N =94; experimental group, N= 169)

	Type of instruction	N	Mean	Std. Deviation	Std. Error Mean
FSWEx (Panning subscale)	Conventional Approach	94	20.5	1.55	0.159
	Inquiry Based Method	169	20.6	1.58	0.122
FSWEx (experimenting subscale)	Conventional Approach	94	21.1	1.49	0.154
	Inquiry Based Method	169	20.8	1.43	0.110
FSWEx (data analysis subscale)	Conventional Approach	94	18.4	1.26	0.130
	Inquiry Based Method	169	18.2	1.27	0.098

Source: Field data (2015).

It was also necessary to analyze and find out whether the mean of the control and experimental groups observed differed statistically or not after the genetics teaching intervention. Hence the SPSS independent samples t-test was used. The FSWEx self-concept subscales pre-test scores for i. planning experiments, ii. practical experimentation and iii. data analysis were analyzed. However, in all three subscales understudy i. planning experiments, ii. practical experimentation and iii. data analysis) no statistically significant difference was found on students self-concept posttest scores based on the type of instruction they received when null hypothesis was subjected to independent samples t-test. Details of t-test findings for all subscales are summarized in table 6.18 (b) below). For example analysis of posttest scores in the planning experiment subscale with t-test for equality of means of IBA (M=20.6, SD= 1.58) and that of TM classes (M=20.5, SD= 1.55); found t (261) =-0.540, p = 0.590, at alpha = 0.05. Hence the null hypothesis that regardless of their groups, Morogoro students did not differ significantly in their sense of ability towards planning science experiments after the intervention was accepted. Independent samples t-test for equality of means for from posttest scores of practical experimentation subscale also failed to reject the null hypothesis at $\alpha = 0.05$. Independent samples t-test for practical experimentation subscale posttest scores of the IBA (M=20.88, SD= 2.96) and that of TM classes (M=21.14, SD= 1.49); produced t (261) =1.429, p = 0.154, at alpha = 0.05. Hence the null hypothesis that there is no statistically significant difference in the sense of ability (self-concept) to conduct practical experimentation between students exposed to inquiry-based teaching (IBA) approach and those exposed to the conventional traditional method (TM) was also accepted.

Lastly, as shown in table 6.18(b), independent samples t-test for equality of variances for posttest scores of analyzing data subscale also failed to reject the null hypothesis at $\alpha = 0.05$. Independent samples t-test for data analysis subscale posttest scores of the IBA (M=18.2, SD= 1.27) and that of TM classes (M=18.44, SD= 1.26); produced t (261) =0.957, p = 0.340, at alpha = 0.05. Hence the null hypothesis that there is no statistically significant difference in the sense of ability to analyze and interpret experimental data between students exposed to inquiry-based teaching (IBA) approach and those exposed to the conventional traditional method (TM) after the intervention was also accepted. These findings imply that the two groups of students that were involved in this quasi-experimental study had a similar sense of the ability in planning experiments, actual experimenting, and data analysis even after eight weeks of genetics intervention. These contradict the initial conception that the students in the experimental group taught using the inquiry-based approach would have higher self-concept towards planning experiments, actual experimenting, and data analysis than the control group taught traditionally.

***6.18 (b)* Independent samples t-test for posttest scores on FSWEx self-concept subscales (n=94 and experimental group n=169)**

		F	Sig.	T	df	Sig. (2-tailed)	Mean Differ	Std. Error Diffe	95% Confi. Interval of the Differ	
									Lower	Upper
FSWEx planning subscale	Equal variances assumed	0.467	0.495	-0.540	261	0.590	-0.109	0.202	-0.50	0.28
	Equal variances not assumed			-0.544	196.34	0.587	-0.109	0.201	-0.50	0.28
FSWEx experimenting subscale	Equal variances assumed	0.094	0.759	1,429	261	0.154	0.267	0.187	-0.10	0.63
	Equal variances not assumed			1.411	185.18	0.160	0.267	0.189	-0.10	0.64
FSWEx data analysis subscale	Equal variances assumed	0.045	0.832	0.957	261	0.340	0.156	0.164	-0.16	0.47
	Equal variances not assumed			0.959	193.85	0.339	0.156	0.163	-0.16	0.47

Source: Field data (2015).

Although it was conceptualized that the experimental group would develop higher positive self-concept towards experimentation than the control group, this was not the case in this study. No statistically significant difference was found on students self-concept posttest scores in all three subscales (i. planning experiments, ii. practical experimentation and iii. data analysis) based on the type of instruction they received the when null hypothesis was subjected to independent samples t-test (table 6.18(b)). These findings differ from those by VeisiKahre et al. (2015) on their quasi-experimental pretest-posttest study with a control group of the first-grade high school male and female students in Holilan in the academic year 2013 to 2014. In this study with multistage cluster sampling, the experimental groups received training in the problem-solving method. The findings showed that problem-solving training causes increased self-concept academic students than the control group. The findings from the current study, however, are in line with the findings by Nath (2015) in his study on constructivist approach as a way of promoting self-concept and achievement in the science of upper primary students. In his study Nath (2015) found that the self-concept of the students was found to be almost the same before and after the treatment. There was no significant difference between self-concept scores of the two groups after the study. After the treatment, the two groups of respondents did not vary statistically in terms of their self-concept (Nath, 2015). This finding signifies that constructivist approach-based experiments as a tool in teaching science did not enhance the self concept of students compared to traditional teaching.

6.3.4.5 Summary of findings on motivation between the control and experimental group

The aim of this part of the study was to compare the effectiveness of inquiry and traditional methods of teaching in developing students' motivation towards science process skills. In this respect, the students' motivation towards science process skills in both the experimental and control groups were scored prior to and after teaching intervention. The control group students and experimental group students were taught differently the topic of genetics in order to compare the level of students' motivational change towards science process skills the end of teaching. The aim was to statistically answer the question, "what are the differences in students' motivation towards science process skills after they experienced inquiry and traditional science instruction?. Preliminary analyses involved testing for violations of assumptions of normality and exploring the descriptive statistics to provide further support for parametric treatment of the data. To verify that the two groups were matched on pretest scores and provide justification for interpreting gain scores for the sample, independent-sample t-tests were performed comparing the inquiry and control group on all pretest measures including science performance.

General linear model with repeated measure for within group effects found a statistically significant within-group effects was noted for a time in all motivation constructs under study (i. intrinsic motivation, ii. grade motivation, iii. career motivation, iv. self-efficacy, v. self-determination, and vi. self-concept). This implies that regardless of the method of teaching in this study, there were significant within-groups effects with regard to the development i. intrinsic motivation, ii. grade motivation, iii. career motivation, iv. self-efficacy, v. self-determination, and vi. self-concept. Statistical significant time effects were noted at alpha =0.05 level. Repeated measure for between - subject effects on all motivation constructs studied accepted the null hypotheses stated. This implies that the between- group interaction effects (method * groups* time) to all constructs were not significant. This means that there were significant gains in the motivation of Morogoro students towards science process skills over time but there was no statistically significant differential improvement among the students of the control group and experimental group over time. An analysis of independent samples t-test based on the type of instruction students received at (α) =0.05 produced a value of p =0.274 for intrinsic motivation, the value of p =0.931 for grade motivation, the value of p =0.598 for career motivation, value of p =0.583 for self-efficacy, the value of p =0.898 for self-determination. An analysis of independent samples t-test based on the type of instruction students received at (α) =0.05 also produced p =0.622 p =0.405 p =0.617 for students self-concepts towards planning experiments, actual experimentation, and data analysis respectively. Hence contrary from the initial conception of the researcher that the experimental group taught genetics through inquiry approach would achieve more on the posttest scores on motivation constructs than the control group. Inferential statistics with t-test as seen in table 6.25 did not prove that. Table 6.19 summarizes all the findings from pretest to posttest of motivation constructs towards science process skills from descriptive to inferential statistics carried out.

Table 6.19 Summary of all the findings from pretest to posttest of motivation constructs understudy

	PRETEST RESULT VALUES				POSTTEST RESULT VALUES				
		Descriptive		t-test	Repeated Measures		Descriptive		t-test
		Mean	s.d		Within F and η^2 values	Betwen F and η^2 values	Mean	s.d	
Intrinsic Motivat	contr	9.92	1.15	t=0.274	F=0.000	F=0.624	18.11	1.45	t=0.569
	exper	9.93	1.18		η2 = 0.949	η2 = 0.001	18.01	1.42	
Self-efficacy	contr	10.6	3.64	t=0.583	F=0.000	F=0.302	15.1	1.45	t=0.340
	exper	10.9	3.91		η2 = 0.678	η2 = 0.01	15.0	1.42	
Grade motv	contr	9.40	1.19	t=0.931	F=0.000	F=0.089	18.4	1.26	t=0.339
	exper	9.39	1.24		η2 = 0.94	η2 = 0.11	18.2	1.27	
self-determination	contr	9.79	0.749	t=0.898	F=0.000	F=0.884	18.6	0.48	t=0.563
	exper	9.81	0.786		η2 = 0.918	η2 = 0.01	18.6	0.47	
Career motv	contr	7.56	2.21	t=0.598	F=0.000	F=0.958	17.4	1.94	t=0.572
	exper	7.41	2.19		η2 = 0.95	η2 = 0.01	17.3	1,89	
Self-concept (plan)	contr	10.6	2.24	t=0.622	F=0.000	F=0.622	20.5	1.55	t=0.590
	exper	10.8	2.30		η2 = 0.831	η2 = 0.01	20.6	1.58	
Self-concept (exper)	contr	9.53	3.13	t=0.405	F=0.000	F=0.428	21.1	1.49	t=0.154
	exper	9.20	2.96		η2 = 0.865	η2 = 0.02	20.8	1.43	
Self-concept (analy)	contr	8.77	3.15	t=0.617	F=0.000	F=0.878	18.4	1.26	t=0.339
	exper	8.57	3.14		η2 = 0.93	η2 = 0.01	18.2	1.27	

Because of the nature of this quasi-experimental study, caution should be taken with generalizing the findings beyond this population of Morogoro students. However, generalization with caution may contribute to the knowledge base of what motivates students within the context of science courses such as science process skills understudy. Regardless of how Morogoro students were taught genetics, after analyzing the conceptual factors that motivate Morogoro students to learn science process skills, the highest motivational construct existed in the perceived self-determination (mean=1.86, sd=0.48 for the control group, mean=18.6, sd=0.47 for the experimental group) and grade motivation(mean=18.4, sd=1.26 for the control group, mean=18.2, sd=1.27 for the experimental group). The higher scores of students in grade motivation, on the other hand, is a reflection of the importance of getting an “A” in their science class. This finding supports the research of Chumbley et al. (2015) who identified those extrinsic motivations such as learning science as a “means to an end” (i.e., competition, rewards, good grades, etc.) can serve as a motivator to learn science Chumbley et al. (2015). On the other hand, regardless of the group, self-efficacy was the lowest motivator for students to succeed in science process skills course (mean=15.1, sd=1.45 for the control group, mean=15.0, sd=1.42 for the experimental group). This means Morogoro teachers are supposed to give students more responsibility and choice during the course of teaching such as providing them with opportunities to plan and evaluate their learning, builds self-confidence and can help maintain high levels of self-efficacy (Zimmerman, 2000). Generally, the results showed that regardless of the method of teaching, students experienced significant increases in levels of all motivation constructs (intrinsic motivation, ii. grade motivation, iii. career motivation, iv. self-efficacy, v. self-determination, and vi. self-concept) participation in the genetics course. Despite the fact that the results did not predict what different benefit students who received inquiry-based teaching with respect to motivations as compared to control group students, however, these results have important implications for designing interventions. A recent trend in educational research highlights the importance of allocating students to different intervention treatment intensities (Barnett et al. 2004) depending on baseline aptitudes..

6.3.4.5 CORRELATION BETWEEN SCIENCE PROCESS SKILLS ACHIEVEMENT WITH COGNITIVE AND MOTIVATIONAL VARIABLES

6.3.4.5. 1 Introduction

Science process skills are known as procedural skills, experimental and investigating science habits of mind or scientific inquiry abilities. The association between science process skills, cognitive abilities, motivational variables and scholastic achievement has been a subject matter of various research studies (Hamilton & Swortzel, 2007; Harlen, 1999; Padilla et al. 1983; Scharmann, 1989). It is claimed that science process skills are needed to better understand the content of science of students (Scharmann, 1989). It was the aim of this study also to determine the existing correlation between student's achievement in science process skills with the achievement in conceptual understanding of contents and motivation. Computer SPSS Pearson's correlation was employed to find out the relationship between student's performance in the science process skills test with his/her conceptual understanding of Biology contents and his/her motivation (intrinsic motivation, self-efficacy, and self-concept) towards science process skills. Pearson product-moment correlation coefficient or Pearson's correlation "p" is a measure of the strength of the linear relationship between two variables. If the relationship between the variables is not linear, then the correlation coefficient does not adequately represent the strength of the relationship between the variables.

Regardless of their groups in the quasi-experimental study, students' posttest scores in the science process skills test (BPST), genetics conceptual test, SMQ II questionnaire, and the FSWEx self-concept scale were analyzed for the variables to be correlated with. This section intends to specifically answer statistically the following questions. Is there a significant correlation between students´ performance in the science process skills;

i. with their conceptual understanding of Biology contents?
ii. with their intrinsic motivation towards science?
iii. with their self-efficacy towards science?
iv. with their self-determination towards science?
v. with their self-concept towards experimentation in science?

The study was also interested in correlating the change of knowledge level of science process skills of students from pretest to posttest of students with their changes in the i. conceptual understanding of genetics, ii. intrinsic motivation towards science iii. self-efficacy towards science iv. self-determination to science, and v. self-concept towards experimentation. The intention was to determine whether the change of science process skills achievement of students as a class correlates with changes in other variables under study. The following questions were answered statistically with the above questions. Is the pretest-posttest change in science process skills of students significantly correlate with their change in their,

i. ***conceptual understanding of Biology contents (genetics being the case)***

ii. in intrinsic motivation towards science?

iii. in self-efficacy towards science?

iv. in self-determination towards science?

v. in self-concept towards experimentation in science?

A Pearson Product Moment Correlation Coefficient (r) was the statistical procedure chosen to ascertain the magnitude of the relationship between the subjects' science process skills and attitudes toward science. This procedure is commonly used in determining the extent of a relationship existing between variables and is probably used most frequently in educational research (Downing et al. 1997). The advantage is that correlational studies do not require large samples. If a relationship exists it is assumed that it will be evident in a sample of moderate size (Ary et al. 1990). In interpreting the size of the correlation coefficient, this study adopted the rule of thumb for interpreting the size of a correlation coefficient as suggested by Hinkle et al. (2003), which is summarized in Table 6.20.

Table 6.20: The rule of thumb for interpreting the size of a correlation coefficient

Size of correlation	Interpretation
.90 to 1.00 (−.90 to −1.00)	Very high positive (negative) correlation
.70 to .90 (−.70 to −.90)	High positive (negative) correlation
.50 to .70 (−.50 to −.70)	Moderate positive (negative) correlation
.30 to .50 (−.30 to −.50)	Low positive (negative) correlation
.00 to .30 (.00 to −.30)	negligible correlation

Hinkle et al. (2003)

6.3.4.5.2 Correlation between science process skills achievement and conceptual understanding of Biology contents (a case study of Genetics)

Educators, who are promoting the inquiry process as the essence of learning science, believe firmly that this approach will make a significant contribution to the conceptual understanding of scientific contents also. They believe also that student who is weak in the content area may not be able to apply these process skills (Harlen, 2000). Computer SPSS Pearson correlation coefficients were computed to determine whether any significant relationship existed between achievements of Morogoro students in science process skills and their achievement in the conceptual understanding of Biology contents. Regardless of their groups in the quasi-experimental study, students' posttest scores in the science process skills test (BPST) and in the genetics, the test was correlated to find out their relationship. As seen in the table 6.21 (a) below a weak positive correlation between students' performance in the science process skills and their performance in the genetics test were seen. This relationship was not significant at 0.005. This implies that science process performance was not proven to be a strong predictor of students' achievement in the conceptual understanding of genetics in this study. This further implies that those students who perfumed well in the science process skills test did not necessarily perform well in genetics test.

Table 6.21(a): Correlation between students' achievement in science process skills and their genetics knowledge

Correlations			
		Posttest scores in BPST	Posttest scores in genetics
Posttest scores in BPST	Pearson Correlation	1	0.045
	Sig. (2-tailed)		0.468
	N	263	263
Posttest scores in genetics	Pearson Correlation	0.045	1
	Sig. (2-tailed)	0.468	
	N	263	263

Source: Field data (2015).

The result of this study showed that mastery of the science process skills does not ensure acquisition of scientific knowledge. The findings that there is no statistically significant relationship between students' performance in the science process skills and their achievement in science contents imply that the teaching of content must take precedent over the training of students on the acquisition of science process skills. In their study, which involved teaching students the science process skills during science experimentation, Padilla et al. (1983) concluded that these complex process skills cannot be learned via a two weeks unit in which science content is typically taught. Rather, experimenting abilities need to be practiced over a period of time. Those having the extended treatment outscored those experiencing the two-week unit. This finding, however, contradicts conclusions made by Millar (1987) who argue that students' ability to use process skills depend on the extent of their knowledge of the contexts they are asked to work on. Studies by other researchers (Song & Black, 1991; Lock, 1993) also found that performance of tasks requiring these process skills is strongly content-dependent.

This study also correlated the change of knowledge level of science process skills of students from pretest to posttest of students with a change in their conceptual understanding of genetics from pretest to posttest. The difference in the BPST scores from pretest to posttest and that of genetics test were used. Table 6.21(b) shows that the pretest-posttest changes in science process skills of students significantly correlate with their pretest-posttest change in their conceptual understanding of genetics. The correlation was significant at the 0.01 level (2-tailed).

Table 6.21(b): Correlation between Students pretest-posttest change in science process skills and their pretest-posttest change in change in Genetics knowledge

Correlations			
		Pretest-posttest change in genetics	*Pretest-posttest change in BPST*
Pretest-posttest change in BPST	Pearson Correlation	1	0.352**
	Sig. (2-tailed)		0.000
	N	263	263
Pretest-posttest change in genetics	Pearson Correlation	0.352**	1
	Sig. (2-tailed)	0.000	
	N	263	263
**. Correlation is significant at the 0.01 level (2-tailed).			

Source: Field data (2015).

A Pearson correlation coefficient value of 0.352 was found between students pretest-posttest change in science process skills of with their pretest-posttest change in genetics knowledge. Table 6.21(b) indicates by the value or r = 0.352 (Hinkle et al. 2003) a moderate positive linear relationship between students pretest-posttest changes in science process skills with their pretest-posttest changes in genetics knowledge which were significant at α =0.01 (2-tailed).

6.3.4.5.3 Correlation between science process skills achievement and the intrinsic motivation of students towards science

Another purpose of this study was to test the hypothesis that there is a positive correlation between students' achievement of science process skills as an independent variable and their intrinsic motivation towards science. Researchers and educators in science education believe that intrinsic motivation is a significantly important factor for academic learning and achievement (Stipek, 1998). It has a positive impact upon learning as it stimulates, sustains and gives directions to an activity. Highly motivated students often require little guidance from the teachers and are capable of doing a higher degree of complicated work independently (Stipek, 1998).

Pearson correlation coefficients for the sum of science process skills posttest scores and for the SMQ-II intrinsic motivation posttest scores were computed to determine whether any significant relationship existed between the two variables. Regardless of their groups in the quasi-experimental study, students' posttest scores in the science process skills test (BPST) and in the SMQ II intrinsic motivation test were correlated to find out their relationship. Table 6.22(a) hints on a moderate positive correlation between students' performance in the science process skills and their intrinsic motivation towards science after genetics intervention were seen. This relationship was not significant at 0.005. This implies that an intrinsic motivation towards science was not statistically proven to be a strong predictor of students' achievement in science process skills this study. This further implies that those students who performed well in the science process skills test did not necessarily demonstrate an intrinsic motivation towards science.

Table 6.22(a): Correlation between students' achievement in science process skills and their intrinsic motivation

Correlations			
		Posttest scores in BPST	Posttest scores in intrinsic motivation
Posttest scores in BPST	Pearson Correlation	1	0.027
	Sig. (2-tailed)		0.658
	N	263	263
Posttest scores in intrinsic motivation	Pearson Correlation	0.027	1
	Sig. (2-tailed)	0.658	
	N	263	263

Source: Field data (2015).

Although the association was not significant, the findings of a weak linear relationship between students' performance in the science process skills and their intrinsic motivation towards science resemble findings from other researchers. Many researchers (Gottfried, 1985; Stipek, 1998; Thompson & Mintzes, 2002) argued that motivation has a significant correlation with cognition, attitude, and acquisition of skills. These findings also resemble the conclusion made by Renni (1990) who found that higher results in science were related to the learner's active engagement in learning tasks, to his or her positive attitude towards the subject and to a highly positive self-concept in science, which all imply the learner's intrinsic motivation to learn.

The study also correlated the change of knowledge level of science process skills of students from pretest to posttest with their change in intrinsic motivation towards science from pretest to post-test. The aim was to find out whether a positive linear correlation exists between changes of students´ science process skills knowledge correlates with their change in intrinsic motivation to science. The resultant difference in students BPST scores from pretest to posttest and that of SMQ II intrinsic motivation test were used. Table 6.22(b) shows that the pretest-posttest change in students' science process skills did not significantly correlate with their change in intrinsic motivation to science. The correlation was not significant at the 0.01 level (2-tailed).

Table 6.22(b): Correlation between students pretest-posttest change in science process skills and their pretest-posttest change in change in intrinsic motivation

Correlations			
		Pretest-posttest change in BPST	*Pretest-posttest change in intrinsic motivation*
Pretest-posttest change in BPST	Pearson correlation	1	0.049
	Sig. (2-tailed)		0.426
	N	263	263
Pretest-posttest change in intrinsic motivation	Pearson correlation	0.049	1
	Sig. (2-tailed)	0.426	
	N	263	263

Source: Field data (2015).

A Pearson correlation coefficient value of 0.052 was found between students' pretest-posttest change in science process skills and their pretest-posttest change in intrinsic motivation to science. According to Hinkle et al. (2003), the value or r = 0.049 indicate almost no linear relationship between the two variables. No linear relationship between students' pretest-posttest change in science process skills and their pretest-posttest change in intrinsic motivation towards science was found. This contradicts the findings by Hough & Piper (1982), who explored the relationship between attitude towards science and science achievement. A significant relationship was found between the pupils' process scores and attitude scores (r = 0.45).

6.3.4.5.4 Correlation between students' science process skills achievement and their self-efficacy towards science

Self-efficacy is a belief in one's abilities to accomplish a task, not a measure of those abilities (Pajares, 2000). It has a positive impact upon learning as it stimulates, sustains and gives directions to an activity. Self-efficacy is the belief that one can succeed in performing particular behaviors; this has been shown to be more strongly related to academic outcomes than many other individual characteristics like student gender, student self-concept, or the perceived usefulness of the knowledge later in the student's life (Pajares and Miller, 1994). Highly self-efficacious students often require little guidance from the teachers and are capable of doing many higher degree of complicated work independently. Hence another purpose of this study was to test the hypothesis that there is a positive correlation between students' achievement of science process skills as an independent variable and their self-efficacy towards science. Researchers and educators in science education believe that self-efficacy is a significantly important factor for academic learning and achievement (Bandura, 1997).

Pearson correlation coefficients for the sum of science process skills posttest scores and for that of SMQ-II self-efficacy posttest scores were computed to determine whether any significant relationship existed between the two variables. Regardless of their groups in the quasi-experimental study, students' posttest scores in the science process skills test (BPST) and in the SMQ-II self-efficacy test were correlated to find out their relationship. As seen in the table 6.23 (a) almost no linear positive correlation between students' performance in the science process skills and their self-efficacy towards science after genetics intervention was found with a Pearson correlation coefficient of +0.04 and the relationship between two variables under study was not significant also at 0.05. This implies that self-efficacy towards science was not statistically proven to be a strong predictor of students' achievement in science process skills this study. This further implies that those students who performed well in the science process skills test did not necessarily demonstrate self-efficacy towards science

Table 6.23(a): Correlation between students' achievement in science process skills and their self-efficacy towards science

Correlations			
		Posttest scores in BPST	Posttest scores in self-efficacy subscale
Posttest scores in BPST	Pearson correlation	1	0.040
	Sig. (2-tailed)		0.519
	N	263	263
Posttest scores in self-efficacy subscale	Pearson correlation	0.040	1
	Sig. (2-tailed)	0.519	
	N	263	263

Source: Field data (2015).

The findings that there is no linear relationship between students' performance in the science process skills and their self-efficacy towards science contradict many findings from other researchers. For example, the findings from the study conducted Downing et al. (1997) to determine if a relationship exists between pre-service elementary teachers ' competency in science process skills and their attitudes toward the field of science. The hypothesis for the study stated that elementary pre-service teachers who demonstrated a high competency in process skills would also indicate positive attitudes toward science. Analysis of the data indicated a significant positive correlation between elementary pre-service teachers' ability to perform science process skills and their attitudes toward science. A recent longitudinal study by Tai et al. (2006) reported repeatedly that students who were otherwise confident and capable believed that they could not learn science, and so they chose not to engage in the activities and experimentation that are the heart of science. Ketelhut (2007) examined the relationship between students' self-efficacy on entry into the authentic scientific activity and the longitudinal data gathering behaviors they employed while engaged in that process. The study found that self-efficacy correlated with the number of data-gathering behaviors in which students initially engaged, with high self-efficacy students engaging in more data gathering than students with low self-efficacy.

This study also correlated the change of knowledge level of science process skills of students from pretest to posttest with their change in self-efficacy towards science from pretest to post-test. The aim was to find out whether a positive linear correlation exists between changes in students as a class in their science process skills knowledge correlates with their class change in their self-efficacy towards science. The resultant difference in students BPST scores from pretest to posttest and that of SMQ-II self-efficacy subscale was used. Table 6.23(b) shows that the pretest-posttest changes in students' science process skills significantly correlate with their change in self-efficacy twards science. The correlation was significant at the 0.01 level (2-tailed).

Table 6.23(b): Correlation between students pretest-posttest change in science process skills and their pretest-posttest change in change in self-efficaccy

Correlations			
		Pretest-posttest change in BPST	*Pretest-posttest change in self-efficacy*
Pretest-posttest change in BPST	Pearson Correlation	1	0.279**
	Sig. (2-tailed)		0.000
	N	263	263
Pretest-posttest change in self-efficacy	Pearson Correlation	0.279**	1
	Sig. (2-tailed)	0.000	
	N	263	263
**. Correlation is significant at the 0.01 level (2-tailed).			

Source: Field data (2015).

As indicated in the table 6.23(b) a moderate positive Pearson correlation coefficient value of 0.052 between students pretest-posttest change in science process skills with their pretest-posttest change in self-efficacy to science was found. However, the correlation was significant at the 0.01 level (2-tailed). According to Hinkle et al. (2003), the value or r = 0.279 indicates a moderate linear positive relationship between the two variables. Hence, the study found a linear relationship between students pretest-posttest change in science process skills with their pretest-posttest changes in self-efficacy towards science.

6.3.4.5.5 Correlation between students' science process skills achievement and their self-concept towards science

The study of self-concept has awakened growing interest in psychological research of recent years. Marsh & Seeshing (1997) defined self-concept as the set of perceptions or reference points that an individual has about himself; the set of characteristics, attributes, qualities and deficiencies, capacities and limits, values and relationships that the individuals know to be descriptive of himself and which he perceives as data concerning his identity. It is the perception that each one has about him or her, formed from experiences and relationships with the environment, where significant people play an important role. The attitudes that one holds toward himself/herself are significantly associated with personal satisfaction and mental health. Thus for a child to achieve, he/she must view himself/herself as able to achieve. In this study, it was also conceived that students' academic self-concept is a good predictor of academic achievement and science experimentation in specific. However, despite the abundance of studies, there are no conclusive studies that clearly identify the direction of the link, which joins these two variables. It has a positive impact upon learning as it stimulates, sustains and gives directions to an activity. Researchers and educators in science education believe that self-efficacy is a significantly important factor for academic learning and achievement (Bandura, 1997). Hence, another purpose of this study was to test the hypothesis that there is a positive correlation between students' achievement of science process skills as an independent variable and their self-concept towards science experimentation.

Students' posttest scores from the three FSWEx Self-concept questionnaire subscales were added to have a general sum which was then correlated with the general posttest sum in the process skills test (BPST). Students' self-concept towards science experimentation was measured by their attitude and self-perception towards planning experiments (06 items), ii. practical experimentation, (06 items) and iii. analyzing experimental data (06 items). A Pearson correlation coefficient for the sum of science process skills posttest scores and for that of FSWEx Self-concept questionnaire subscales posttest scores were computed to determine whether any significant relationship existed between the two variables. Regardless of their groups in the quasi-experimental study, students' posttest scores in the science process skills test (BPST) and in the FSWEx Self-concept questionnaire were correlated to find out their relationship. Table 6.24 (a) shows a weak linear positive correlation between students' performance in the science process skills and their self-concept towards experimentation in science after genetics intervention. A Pearson correlation coefficient of 0.053 was found. And the relationship between two variables under this study was not significant also at 0.05. This implies that self-efficacy towards science was not statistically proven to be a strong predictor of students' achievement in science process skills this study. This further implies that those students who performed well in the science process skills test were not necessarily demonstrating self-concept towards science experimentation.

Table 6.24(a): Correlation between students' achievement in science process skills and their self-concept towards experiments in science

Correlations			
		Posttest scores in self-concept scale	Posttest scores in BPST
Posttest scores in self-concept scale	Pearson Correlation	1	0.053
	Sig. (2-tailed)		0.390
	N	263	263
Posttest scores in BPST	Pearson Correlation	0.053	1
	Sig. (2-tailed)	0.390	
	N	263	263

Source: Field data (2015).

The findings that there is no linear relationship between students' performance in the science process skills and their self-concept towards science experimentation contradict many findings from other researchers. Some previous research works suggest that there is a positive relationship between academic self-concept and academic achievement as measured by grade point average (Cooley, 2000; Gerardi, 2009). For example, the recent study conducted by Jansen et al. (2014) on academic self-concept in science, multidimensionality, relations to achievement measures, and gender differences concluded that students' academic self-concept is a good predictor of academic achievement. In this study, results indicate that the relations between the self-concept and achievement are substantial and subject-specific when grades are used as achievement indicators and that female students possess a lower self-concept in science. The purpose of the study by Campbell & Martinez-Perez (1976) was to test the hypothesis that there are positive correlations among (1) attitudes toward science, (2) self-concept, and (3) achievement of science process skills. In this study, data were analyzed using Pearson's product-moment correlation coefficients. There were positive correlations between: (1) basic science process skills and integrated science process skills, (2) basic science process skills and attitudes toward science, (3) basic science process skills and self-concept, (4) integrated science process skills and attitudes toward science, (5) integrated process skills and self-concept, and (6) attitudes toward science and self-concept.

However, the findings from this study resemble the findings by Afuwape (2011) study that investigated the relationship between students' self-concept and their academic performance in Basic Science. The result showed that there was no significant relationship between the secondary school students' self-concept and their academic performance in Basic Science. It also showed that there was no significant difference between the self-concept of male and female students in Basic Science as well as their performances. In the same line Hoge et al. (2012) conducted a two-year longitudinal study of 322 sixth and seventh graders in which they compared three levels of self-concept (high, middle and low) and then studied the effects of self-concept on achievement and achievement on self-concept. They found that the influence of self-concept on grades was weak but grades had a modest influence on subsequent discipline-specific self-concept. The researchers concluded that past correlation studies have overstated the influence of self-concept on grades and of grades on self-concept.

This study also correlated the change of knowledge level of science process skills of students from pretest to posttest with their change in self-concept towards science from pretest to post-test. The aim was to find out whether a positive linear correlation exists between changes in students` science process skills knowledge correlates with their class change in their self-efficacy towards science. The resultant difference in students FSWEx self-concept questionnaire scores from pretest to posttest and that of science process skills test (BPST) were used. Table 6.24 (b) shows that the pretest-posttest changes in students' science process skills did not significantly and linearly correlate with their change in self-concept to science experimentation. Pearson correlation coefficient value was 0.089 and the correlation was not significant at the 0.01 level (2-tailed).

Table 6.24(b): Correlation between students pretest-posttest change in science process skills and their pretest-posttest change in change in self-concept

Correlations			
		Pretest-posttest change in BPST	*Pretest-posttest change in self-concept*
Pretest-posttest change in BPST	Pearson Correlation	1	0.089
	Sig. (2-tailed)		0.149
	N	263	263
Pretest-posttest change in self-concept	Pearson Correlation	0.089	1
	Sig. (2-tailed)	0.149	
	N	263	263

Source: Field data (2015).

An insignificant positive Pearson correlation coefficient value of 0.089 between students pretest-posttest change in science process skills with their pretest-posttest change in self-efficacy to science was found. According to Hinkle et al. (2003), the value or r = 0.89 indicates the absence of a linear positive relationship between the two variables. Hence the study did not found a linear relationship between students pretest-posttest change in science process skills with their pretest-posttest changes in their self-concept towards science experimentation.

6.3.4.5.6 Correlation between science process skills achievement and the Self-determination of students towards science

Self-determination is the process of utilizing one's will. According to Pintrich & Schunk (2002) self-determination requires that people accept their strengths and limitations, be cognizant of forces acting on them, make choices, and determine ways to satisfy needs. Both the intrinsic motivation and extrinsic motivation are encompassed by the self-determination theory. Studies on self-determination theory indicate that intrinsic motivation (doing something because it is inherently interesting or enjoyable), and thus higher quality learning, flourishes in contexts that satisfy human needs for competence, autonomy, and relatedness (Ryan & Deci, 2000). Students experience autonomy when they feel supported to explore, take initiative and develop and implement solutions for their problems. Students experience relatedness when they perceive others listening and responding to them. When these three needs are met, students are more intrinsically motivated and actively engaged in their learning (Ryan & Deci, 2000).

Another purpose of this study was to test the hypothesis that there is a positive correlation between students' achievement of science process skills as an independent variable and their self-determination towards science. Researchers and educators in science education believe that students' self-determination is a significantly important factor for academic learning and achievement (Glynn & Koballa, 2006; Ryan & Deci, 2000). It has a positive impact upon learning as it stimulates, sustains and gives directions to an activity. Highly self-determined students often require little guidance from the teachers and are capable of doing many higher degree of complicated work independently (Stipek, 1998).

A Pearson correlation coefficient for the sum of science process skills posttest scores and for that of SMQ II self-determination posttest scores were computed and subjected to SPSS to determine whether any significant relationship existed between the two variables. Regardless of their groups in the quasi-experimental study, students' posttest scores in the science process skills test (BPST) and in the SMQ II self- determination test were correlated to find out their relationship. A moderate positive ($r = 0.34$) correlation between students' performance in the science process skills and their self-determination towards science after genetics intervention was seen (table 6.25a). This relationship was not significant at 0.005. This implies that self-determination of students towards science was not statistically proven to be a strong predictor of their achievement in science process skills in this study. This further implies that there were those students who demonstrated higher self-determination towards science somehow performed well in the science process skills test and vice versa.

6.25 (a) Correlation between students achievement in science process skills and their self-determination towards experiments in science

Correlations			
		Posttest scores in BPST	Posttest scores in self-determination scale
Posttest scores in BPST	Pearson Correlation	1	0.34
	Sig. (2-tailed)		0.492
	N	263	263
Posttest scores in self-determination scale	Pearson Correlation	0.34	1
	Sig. (2-tailed)	0.492	
	N	263	263

Source: Field data (2015).

Although the association was not significant, the findings that there is a moderate positive linear relationship between students' performance in the science process skills and their self-determination towards science resembles findings from other researchers. The positive and significant relationships, to a certain extent, can be considered meaningful and taken as evidence for possible causal relationships between these variables. Science education researchers (Glynn et al. 2009; Glynn et al. 2011; Chow & Yong, 2013; Glynn & Koballa, 2006; Garcia & Pintrich, 1996) found that higher results in science were significantly related to the learner's positive attitude and the self-determination towards the subject. For example, a study conducted by Chow & Yong (2013) to investigate students' motivation and achievement in combined science demonstrated significant differences in motivational orientations towards learning-combined science between boys and girls while the correlation analyses showed that there were significant positive associations between students' self-determination and science achievement. Garcia & Pintrich (1996) on the other hand found that the self-determination and intrinsic motivation of college biology students increased when the students could select the course readings and term paper topics as well as the due dates for class assignments. Painter (2011) on the other hand conducted a study to examine a proposed motivational model of science achievement based on the self-determination theory. His study relied on the United States eighth-grade science data from the 2007 third international mathematics and science study to examine a structural model that hypothesized how perceived autonomy support, perceived competence in science, intrinsic motivation, and science achievement related to each other. The study found that the strongest direct effect on science achievement was students' perceived competence and self-determination in science. Autonomy support had positive direct effects on students' perceived competence in science and intrinsic motivation and had indirect positive effects on science achievement (Painter, 2011).

The study also correlated the change of knowledge level of science process skills of students from pretest to posttest with their change in self-determination towards science from pretest to post-test. The aim was to find out whether a positive linear correlation exists between changes in students as a class in their science process skills knowledge correlates with their class change in self-determination to science. The resultant difference in students BPST scores from pretest to posttest and that of SMQ-II self-determination test were used. Table 6.25(b) shows that the pretest-posttest changes in students' science process skills did not significantly correlate with their change in self-determination to science. The correlation was significant at the 0.01 level (2-tailed).

Table 6.25(b): Correlation between students pretest-posttest change in science process skills and their pretest-posttest change in change in self-concept

Correlations			
		Pretest-posttest change in BPST	*Pretest-posttest change in self-determination*
Pretest-posttest change in BPST	Pearson Correlation	1	0.244**
	Sig. (2-tailed)		0.000
	N	263	263
Pretest-posttest change in self-determination	Pearson Correlation	0.244**	1
	Sig. (2-tailed)	0.000	
	N	263	263
**. Correlation is significant at the 0.01 level (2-tailed).			

Source: Field data (2015).

A Pearson correlation coefficient value of 0.244 was found between students pretest-posttest change in science process skills with their pretest-posttest change in self-determination to science. According to Hinkle et al. (2003) the value or r = 0.244 indicates a moderate positive linear relationship between the two variables. Hence, this study found a significant linear association between students pretest-posttest change in science process skills with their pretest-posttest change in self-determination towards science.

CHAPTER SEVEN

COMPARING THE PERFORMANCE OF MOROGORO BIOLOGY STUDENTS IN THE SCIENCE PROCESS SKILLS TEST AT THE SECOND AND THIRD STAGES OF THE STUDY

7.1 Introduction

This chapter compares the performance of Morogoro Biology students in the science process skills test at the second phase of the study and during the third stage. As it has already been elaborated earlier, at the second stage of this research, the developed test BPST was employed to examine the knowledge level of science process skills of the advanced level secondary school Biology students in Morogoro. The study also assessed the performance of students in each of the five process skills (formulating hypotheses, defining operationally, identifying and controlling variables, design experiments as well as and interpreting data) understudy in this phase. In the third stage, the study assessed the effectiveness of the inquiry-based approach to the development of students' science process skills, content understanding and motivation by comparing it with the conventional approach. This part discusses and compares the general performance of Morogoro students during these two phases of the study. It also compares the performance of these students on individual science process skills understudy in these two phases of the study.

However, it has to be noted that the sampling frame and sample size of students who participated in these two phases were different. For example, in the second phase, the test was administered to a group of 353 advanced level Biology students from all four Biology schools present in the municipality. Of the 353 students in this phase, 246 (69.7%) were female and 107(30.3%) were male students. In the third stage, the study had 94 (35.7%) control group students who were taught themes in genetics using the conventional method and 169(64.3%) experimental group students who were taught genetics using inquiry-based approach (IBA). Only three schools (Kilakala, Alfa germs, and Bigwa sisters) were involved in the third phase as opposed to 04 schools in the second phase. On the other hand, one cohort which took part in the pilot and second stages had already completed their secondary education and hence they were not available to participate in the third phase. The new cohort of the incoming Biology students had to replace them. However, the comparison of the performance between these phases is possible because in both cases, the study employed the test that was developed and validated in the first stage, BPST in examining the knowledge level of science process skills. The chapter is sub-divided into two main section. Section 7.2 compares the general performance of students in BPST in two phases of the study and section 7.3 compares the performance of students in each of the five specific process skills under study.

7.2 Comparing the general performance in two phases of the study

In the first assessment when this study was in its second stage, descriptive statistics indicated the mean score of 353 students who participated in the study to be 17.2 (49.1%) with s.d of 7.3. The highest score was 28 (80%) and the lowest 09 (25.7%) out of 35 possible. The overall mean score was 17.2 (49.1%) which means that on average, the advanced level Biology students in Morogoro scored between 17 to 18 items correctly out of 35 total questions. According to the grading system of Tanzania adopted in this study, 49.1% represents a "C" class which means average knowledge level. This further means that, on overall, Morogoro Biology students have barely average knowledge level of science process skills when the BPST was applied for the first time. Hence, much work has to be done if the objectives of competence based curriculum have to be realized. Table 7.1 A below summarizes comparison of students' performance in process skills test BPST at first assessment, during pretest, and at posttest.

Table 7.1 A comparison of students' performance in the science process skills test BPST at first assessment, during pretest, and at posttest

1st Assessment		***2nd Assessment***				
BPST (n=353)		Pretest BPST (n=263)			Posttest BPST (n=263)	
Mean	Sta. dev		Mean	Sta. dev	Mean	Sta. dev
17.2 (49.1%)	7.3	Control	15.2 (43.4%)	2.84	19.1(54.5%)	1.97
		Experim	15.4 (44.0%)	2.44	19.7 (56.2%)	1.82

Research data, 2014 & 2015

Table 7.1 above summarize the shift in the general performance of Morogoro students from the first to the second assessment. In the second assessment (third stage), the mean of pretest scores in the BPST of the students in the control group was 15.2 out of 35 while the mean for the students in the experimental group was 15.4. The standard deviation (a spread of individual scores around their respective means) were 2.84 for the control group and 2.44 for the experimental group. On posttest, the mean of students score in the experimental group was 19.7 out of 35, while the mean of the control group 19.1 out of 35 items. The spread (standard deviation) of individual scores around their respective means changed from 2.84 to 1.82 for the control group and from 2.44 to 1.97 for the experimental group (see table 7.1 above).

The relative higher performance of students at posttest in the second assessment (third stage) as seen in table 7.1 above can directly be linked as a result of genetics course intervention. As it can be seen in table 7.1, the mean of the experimental group rose from 44% at pretest to 56.2% after intervention while that of the control group rose from 43.4 to 54.5% at posttest. This means that the performance at posttest was relatively good in both groups as compared to the performance in the first assessment (second phase) with the overall percentage mean of 49.1% and s.d of 7.3 (see table 7.1 above). The main reason for the shift in the performance might be the genetics intervention using inquiry and conventional approaches. This argument is supported by Germann et al. (1996) who asserted that students' good performance on science process skills was dependent on their experience with and domain-specific practice activities on the skills in prior tasks. Ruiz-Primo & Shavelson (1996) also reported that student scores depended on the particular tasks investigated and on the particular method used to assess their performance. In the similar vein, Millar & Driver (1987) found that students' ability to use process skills depend on the extent of their knowledge of the contexts they are asked to work on. This is also explained by the finding (Rowe & Foulds, 1996; Tobin & Capie, 1982) that performance of tasks requiring these process skills is strongly content-dependent.

However, this shift in performance by Morogoro students in the science process skills test after intervention as seen in the second assessment still cannot be regarded as "good" considering the high premium placed on the acquisition of science process skills in their science curricula. Some of the possible reasons for the students' weak performance might be that many might not be familiar with the types of tasks investigated and assessment used in this study. The current study holds the view that eight weeks of genetics intervention in this study were not enough for students to be able to acquire competence and mastery a full range of scientific skills. Akcay & Yager (2015) also holds the same view that achieving scientific literacy takes time, and call for dramatic changes in what students are taught, in how student performances are assessed, in how teachers are educated and stay current, and the complex relationships that exist between school and community. Nevertheless, the intervention in this study especially that of using inquiry-based approach has produced a significant positive result in the development of students science process skills development.

The lesson here is that Tanzania science teachers need to capitalize on opportunities in the activities done in the science classroom to emphasize science process skills. Science process skills exercised in relation to some science content, and have a crucial role in the development of learning with understanding (Harlen, 1999). There is a problem of how to integrate content and process of science in Tanzania. Osaki (2007) attributed this to poor science teacher preparation in teacher training institutions. According to the author, teacher education curriculum has failed to promote reflective practices and constructivist approaches to prospective science teachers. As a result, these institutions are increasingly producing teachers who are weak in practical skills especially laboratory experiences (Osaki, 2007). A number of publications (e.g. Tilya, 2003; Mafumiko, 2006; Osaki 2007) show that the predominance of talk and chalk instead of learner-centered/active approach dominates science teaching and learning. Theoretical teaching which has been noticed to dominate sciences hardly help students to develop required knowledge, skills, and attitude.

7.3 A comparison of students' performance by specific science process skills between first assessment and second assessment

Another aim of this chapter is to make a comparison of students' performance by specific science process skills between first assessment and second assessment. In the first assessment when this study was in its second stage, the findings with BPST indicated that, correct response percentage for Morogoro students was highest for the process skills of identifying and controlling variables with the mean score of 4.05(57.7%) and was lowest for the process skill of data interpretation with the mean score of 2.34(33.4%). Surprisingly however in the third stage after implementation of genetics course using different approaches to teaching, correct response percentage for Morogoro students was highest for the process skills of formulating hypothesis with a mean of 4.18 (59.7%) for the control group and mean of 4.21 (60.1%) for the experimental group. This finding might imply that Morogoro students had more opportunities to practice the skill of stating and formulating hypotheses during genetics intervention than the opportunities for identifying and controlling variables. Mastery of the hypothesis skill according to Darus and Saat (2014) includes understanding of the concept of hypothesis and variable, understanding of the effect one variable has over another, stating the relationships between variables, formulating hypotheses, differentiating a hypothetical statement from a non-hypothetical statement, and determining whether a hypothesis can be tested to ascertain its plausibility (see table 7.2 below).

However, the findings that Morogoro Biology students have performed relatively better in items measuring hypothesis formulation skills after intervention were surprising. This is because Many researchers who have studied hypotheses formulation within science education have concluded that students still have weak abilities in formulating and testing hypotheses. For example, in their study on young children differentiation of hypothetical beliefs from evidence, Sodian et al. (1991) found that students tend to produce or repeat the effect rather than to discover its causes and they have trouble on identifying likely causes. Nevertheless, the finding of this study is the highlight that better preparation of students for the future may require new teaching approaches that respond to and focus on not only learning scientific content but also on acquiring advanced transferable abilities such as the ability to formulate testable and plausible hypotheses. As the observation made by Filson (2001) that students have difficulty with hypothesis because their books and lessons mention hypothesis, but almost never really explain or model them and frequently hypotheses are confused with theories.

Table 7.2 A comparison of students' performance in the science process skills test at first assessment, during pretest and at posttest

Scientific skill	***1st Assessment (second study phase)***		***2nd Assessment (third stage of the study)***				
	BPST (n=353)		Pretest BPST (n=263)			Posttest BPST (n=263)	
	Mean	Sta. dev	Mean		Sta. dev	Mean	Sta. dev
Hypothesis formulation	3.49	1.43	Control	4.18	1.19	4.18	1.19
			Experim	4.21	1.31	4.41	1.31
Controlling variables	4.05	0.88	Control	2.63	2.25	3.02	2.61
			Experim	2.78	2.11	3.00	2.51
Design experiments	3.27	0.96	Control	3.03	0.92	3.13	0.92
			Experim	3.14	0.84	3.54	0.84
Data analysis & interpret	2.34	0.75	Control	2.24	0.74	2.37	0.82
			Experim	2.36	0.66	2.75	0.66
Defining operationally	3.71	0.96	Control	2.67	0.80	3.74	0.97
			Experim	2.80	0.63	3.67	0.94

Source: Research data (2015)

The inconsistency that Morogoro students scored better in the skill of controlling variables in the second stage while in the third stage students scored better in the skill of formulating hypothesis might be attributed to two facts. Firstly, the difference in the sample size between these stages of the study. It has to be noted that the number of Morogoro Biology students who took BPST in the second stage were 353 from four (04) different schools while the number of students who participated quasi-experimental intervention (the third stage) was 263 only from three schools. Due to time constraints of the researcher, and difficult school administrative arrangements, one potential school with a total of 90 students (44 form V, and 46 form VI) with the name of Lutheran Junior Seminary did not participate in the quasi-experimental study. The ninety (90) students who did not undertake BPST in the intervention phase might have caused this shift of scores. Secondly, the shift in performance might also be caused by the effect of 8-week genetics teaching. Intervention might have improved students' knowledge level of hypothesis more than how it improved their ability in controlling experimental variables. The mean of a raw score of the subtest was low, indicating that the students found the subtest more difficult (see table, 7.2 above).

As it was in the first assessment in the second stage, the mean score on BPST of Morogoro students in the second assessment was also lowest for the process skills of data analysis and interpretation. Table 7.2 above shows that students had a mean score of 2.37(33.8%) for the control group and the mean of 2.75 (39.2%) for the experimental group students even after posttest. In the first assessment, analysis of BPST scores indicated that students' mean score was only 2.34(33.4%) out of seven (07) items in this subscale. These findings that Morogoro students had poor scores in data analysis even after posttest resembles the findings reported by both Hamilton & Swortzel (2007) and Dyer et al. (2004) where students scored higher on questions measuring their skills of controlling variables but scored poorly on items measuring their ability in graphing and data interpretation. These finding on Morogoro students also correspond with those by Hackling & Garnett (1991) who conducted a research on students ability in carrying out experiments and found that students at all levels showed a poorly developed skill of problem analysis, planning, and carrying out controlled experiments. Another similar finding is that by Foulds & Rowe (1996) who found that students were capable of identifying all variables influencing an experiment, scoring about 50% on the test items and they could also produce testable hypotheses, with scores of about 40%. However, they were unable to design a controlled experiment and analyze experiment results, gaining an average mark of only 18%. The complexities surrounding understanding of the concept of data analysis extend to science teachers. In an early study, Shadmi (1981) studied science teachers' understanding of the control of variables and found that most teachers had difficulty interpreting the results in the context of experimental settings. The students 'poor performance on the skills of analyzing and interpreting data might be due to the likelihood that they had not been taught well enable them to handle this skill. It is in the view of this study that teacher-centered model of teaching science in the sampled schools in Morogoro, which did not allow the students to practice and internalize the skills over a fairly long period, was likely to be one of the main reasons for the students' poor performance on the skills. This means that current teaching-learning processes should not only focus on conceptual understanding of science, but it must also move in directions similar to those identified in science education research as ‘doing science’ and ‘knowing about science’ (Zimmerman, 2007). Nevertheless, the findings from this subscale pinpoint that the eight weeks of genetics intervention in this study were not enough for students to be able to acquire competence and mastery in analyzing and interpreting data.

The performance of students in the skill of controlling variables dropped from an average of 4.05 (57.8%) in the first assessment to 3.02 (43.1%) for the control group and 3.00 (42.8%) for the experimental group after posttest. This means Morogoro students had below average even after genetics course in the items measuring their ability in controlling variables. This study holds the view that the eight weeks of genetics intervention were not enough for students to be able to acquire competence and mastery in controlling variables. It has to be noted that, across many studies, it is evident that most students and even some adults do not have a generalized understanding of controlling variables because of their ability to identify, select, or design controlled experiments depends on the task content or situational factors (Koslowski, 1996; Linn et al. 1983; Zimmerman, 2000). This skill provides students with the scope and understanding needed to carry out controlled and reliable experiments that might eventually lead to trusted outcomes and valid inferences (Chen and Klahr, 1999). The findings from this study are in congruence with the finding obtained in the study by Turaib (2015) in his study assessing students’ understanding of the control of variables across three grade levels and gender in the United Arab Emirates (UAE). His findings revealed that students across grade levels exhibited alternative conceptions of key ideas related to control of variables. In the related findings, Boudreaux et al. (2008) found that although most of the students participating in their study were able to realize the importance of having controlled conditions for experimentation, many students had difficulties in providing a valid justification for why controlled conditions were important. Research studies in this area call for critical investigations to suggest and develop methods and approaches needed to help students develop sound and coherent understanding of this crucial and essential skill (Zimmerman, 2000). The findings with Morogoro students highlights the need for teachers to pay attention to the development of argumentation and analytical skills needed to argue for which variables need to be manipulated and which ones need to be kept constant.

Analysis of students score in the subscale of designing experiments indicated that the performance in the first assessment was almost the same with that of the second assessment. In the first assessment, students had a mean score of 3.27 (46.7%) out of (07) seven items. In the second assessment, the control group of had a mean of 3.13 (44.7%) while that of the experimental group was 3.54 (50.5%) in this subscale. This means that most of the Biology students still had below average scores in designing experiments even after genetics intervention. The findings that Morogoro Biology students have below average performance in items measuring ability in designing experiments in all assessments were also not surprising. This study holds the same view that the eight weeks of genetics intervention were not enough for students to be able to acquire competence and mastery in designing experiments. A number of science education researchers (Coil et al. 2010; Chen & Klahr, 1999; Adey & Shayer, 1990; Ghanem, 2003) attribute poor students' ability in correctly designing experiments to misconceptions and inaccuracies regarding randomization, sample size, and inability to identify and control variables and poor stated hypotheses. According to Adey & Shayer (1990) students weak in designing experiments because they are rarely given an opportunity to think deeply about experimental design or asked to develop experimental protocols on their own. Scores from BPST showed that most of Morogoro students know that an experiment should contain a control, but many find it difficult to define exactly what a control is. Similar observation was made by Klymkowsky et al. (2011) in their study which intended to reveal student thinking about experimental design and the roles of control experiments. In this study Klymkowsky et al. (2011) surprisingly found that a high percentage of students had difficulty identifying control experiments even after completing three university-level laboratory courses. Some students confused positive control experiments with experiments or negative control experiments. To address this problem Klymkowsky et al. (2011) designed and ran a revised cell biology lab course in which students participated in a weekly experimental control exercise. Not unexpectedly, the results indicate that the revised course led to greater improvements in students' ability to identify and explain the purpose of control experiments. So it can be concluded that using a simple experimental measure, students can become engaged in the process of scientific inquiry, and in turn, begin to think deeply about experimental design.

CHAPTER EIGHT

CONCLUSION AND RECOMMENDATIONS FROM THE STUDY

8.1 Introduction

This chapter presents some conclusions from the study and the recommendations arising from the findings in each phase. The chapter also highlights some limitations and constraints arisen during the study and recommends some areas for further research. The study was conducted in order to establish a base level of information on the teaching and assessment of science process skills in the context of the 2005 competence-based curriculum in Tanzania. Science process skills are one of the competencies heavily advocated in the 2005 competence based curriculum. As it is been elaborated in the previous sections, this study was conducted stage wise in three phases. The chapter is divided into three sections according to the stage of the study. In each stage, a summary of findings and conclusion are provided followed by some recommendations to teachers, researchers, and the Ministry at large. The last part of each section is a summary of some constraints and limitations to the specific study while the last part provides areas recommended for further research.

8.2 Development and validation of a test for measuring students' knowledge level of science process skills (BPST)

8.2.1 Summary of findings and conclusion from phase one of the study

Analysis of pilot study data gathered from the developed test (BPST) indicated that the test is reliable and valid enough to be employed in the large scale. The developed test is sought to be used as an alternative tool in assessing the knowledge level of science process skills of students in secondary education level. Psychometric qualities of the developed BPST to some extent are in congruence with the qualities of many science process skills tests published. Through careful attention to the standards for developing validity arguments of a psychometric test, the study has provided comparative validity evidence related to testing content, response process, and internal structure. The results of the iterative process of item construction, administration, and revision provide support that the BPST with the underlying conceptualization of scientific literacy that sought to be assessed. In addition, the development of this measure was guided by experts from the university of Wuppertal and the ministry of Education in Tanzania with the sound conceptualization of scientific literacy based on the extant literature. In summary, the developed Biology process skills test (BPST) had internal consistency reliability of 0.80 Cronbach alpha, a difficulty index of 0.447 and an overall discrimination index of 0.48. Furthermore, the content validity of BPST is 0.88 concurrent validity of 0.51 and a construct validity (discriminant correlation coefficient) of 0.34. Suitable readability level of BPST is grade 12.

The BPST items in demonstrated good reliability and the items on each adheres to recommended guidelines for percent correct, discrimination index, item-total correlation coefficients, and frequency distribution of distracters selected, all of which provide evidence for the strong internal structure of this measure. For these reasons, we recommend the use of this measure with middle school students and encourage users to evaluate the reliability in their data and consider the appropriateness of this tool for providing valid evaluations of scientific literacy in the contexts in which it is used. The test specifically measured students' performance on specific integrated science process skills which includes skills in i. formulating hypotheses, ii. defining variable operationally, iii. identifying and controlling variables, iv. planning investigations as well as v. analyzing and interpreting data. The test consists of 35 multiple-choice items and took into account realities of the Tanzania education system. It was a multiple choice, four-option format with an emphasis on the use of pictures and drawings to clarify and enhance items. The first version of the 35-item test of Process Skills in Science was validated and administered to 610 eleventh-, twelfth-, and thirteenth-grade students to establish test reliability and to compute item difficulty and discrimination indices to aid in test revision. Evidence of content validity, construct validity, internal consistency reliability, difficult index, and discrimination index was investigated to prove its psychometric properties.

8.2.2 Recommendations to science teachers and researchers in Tanzania with respect to the use of the developed BPST and test construction

i. Assessment is an integral part of instruction, as it determines whether or not the goals of education are being met or not. Assessment affects decisions about grades, placement, advancement, instructional needs, curriculum, and, in some cases, funding. With the implementation of the competence-based curriculum of 2005 in Tanzania and its emphasis on science process skills, Tanzania teachers need to improve their ways of assessing students in these skills. They should strive to assess not only science contents but also the processes of science.

ii. The issue of validity and reliability of classroom-based achievement tests in Tanzania secondary schools have not engaged the attention of many teachers. Designing tests is an important part of assessing students understanding of content and their level of competence in applying what they are learning. A well-designed test can be a useful learning tool for students; a poorly designed test, by contrast, can create a frustrating experience that only assesses students' abilities to take it. Tanzania teachers need to improve their test construction skills especial the skills of measuring complex learning outcomes such as science process skills so that they have the capacity of constructing tests that have psychometric qualities of validity and reliability.

iii. For students to demonstrate the acquisition of integrated process skills, assessment using hands-on procedures (authentic assessment) is pivotal. The use of a paper and pencil test to assess practical skills like BPST has been criticized by several researchers who advocate for practical manipulation of apparatus and physical demonstration of practical skills. Therefore teachers are recommended to use the developed tool (BPST) not as an alternative but an integral tool with authentic measuring procedures.

iv. The researcher recommends teachers to use the developed instrument to get prompt feedback on their competence in science process skills so that they are able to identify areas where they may need remediation. This is because its format does not require expensive resources, and it can easily be administered to large groups of learners at the same time, so it is cost effective and convenient.

v. The developed instrument could be readily adapted to local use to monitor the acquisition of science process skills by the learners. It is recommended that results from the test developed might be used in order to establish a base level of information on whether or not Tanzania students are acquiring the science process skills as emphasized by the competence based curriculum in Tanzania. The results of which could feedback on the effectiveness of the implementation of science curriculum emphasizing the acquisition of science process skills.

vi. The issue of validity and reliability of classroom-based achievement tests in Tanzania secondary schools has not engaged the attention of many science education researchers. Considering the fact that assessment drives student learning, the researcher of the current study recommends other science education researchers in Tanzania to use the procedure used to develop the test instrument as a model for the development and validation of other similar assessment instruments so that enough pool of tests are available.

vii. The researcher of the present study recommends to other science educators and researchers to use the developed instrument in various ways. For instance, researchers who need a valid and reliable instrument to work with, may use the test to; identify the process skills inherent in certain curricula material, determine the level of acquisition of science process skills in a particular unit; establish science process skills competence by science teachers, or to compare the efficacy of different teaching methods in imparting science process skills to learners.

8.2.3 Recommendations to the Ministry of Education and Vocational Training of Tanzania

The following are recommended to the ministry of education and vocational training which is responsible for the formulation, monitoring, and evaluation of the implementation education policies in Tanzania. These recommendations are as a result of the findings in the first stage of this research;

i. The search of available integrated science process skills tests showed the continued need and relative scarcity of a test geared to secondary school students and associated with any particular science curriculum. However, it has been observed that most of Tanzania science teachers lack science process skills test construction skills thus there a scarcity of these tests in Tanzania. The researcher recommends that the Tanzania Ministry of Education has to establish science teacher's capacity building programs in science process skills test construction. One of the major reasons as to why Tanzania science teachers opt to measure students achievements in science contents only instead of integrating both contents and scientific skills have been inadequate test construction skills of scientific process test.

ii. The study also recommends to the Ministry of Education of Tanzania to design pre-service science teachers curriculum with the component which emphasizes on science process skills test construction skills in both ways authentically and using objective tests. The curriculum for pre-service teachers in the training program should involve learning how to develop scientific process skills tests, how to practice using standardized assessment instruments, and to how to interpret the results from testing. The pre-service science teachers need to be prepared in the area of test construction and validation in order to be able to develop valid and reliable tests that will yield accurate feedback of students' science process skills. Test construction skills include the competencies needed for developing quality tests based on stipulated principles of test construction.

8.2.4 Limitations of the study

This part summarizes the constraints and limitations occurred during development and validation of a test for measuring students' knowledge level of integrated science process (BPST). The study has certain limitations that should be taken into consideration when interpreting the results. The limitations are as follows;

i. The BPST constructed seem to be a reasonably good measure of integrated science process skills which are necessary for conducting scientific investigations. However this psychometric study was conducted in only one point of the country, Morogoro municipality using 610 Biology students from 07 selected schools. More proof from bigger samples should have been collected from different parts of the country to establish its relevance and quality.

ii. For students to demonstrate the integrated process skills, assessment using hands-on procedures (authentic assessment) to determine skill acquisition by students is pivotal. The use of a paper and pencil test to assess practical skills like BPST has been criticized by several researchers who advocate for practical manipulation of apparatus and physical demonstration of practical skills. This presents a limitation in the sense that the instrument developed in this study does to accommodate these requirements.

iii. In this study, the test of integrated science process skills (TIPS-II) which was developed by Burns et al. (1985) was used in the determination of concurrent validity or concurrent validity of the developed Biology process skills test (BPST) through comparison of student scores. TIPS-II, however, has some constraints which could have led to the learners' poor performance on it. Comparison of performance of learners on the developed test with their performance in any other alternative locally developed assessment instrument could perhaps have been a better criterion to use in determining the external or concurrent validity of the developed test instrument. However, literature search provided no single test of science process skill that has been designed for Tanzania students, hence there was no alternative.

8.2.5 Areas for further research

The results from this study on the development and validation of BPST present several further research opportunities in the following areas;

i. This study on the development and validation of a tool to be used in the assessment of students' science process skills opens an opportunity for another study to develop and validate an instrument that may be used to determine Tanzania teachers' competence in integrated science process skills. The information about teachers' knowledge level of science process skills is crucial because the competence of students in this area depends on the competence of their teachers also.

ii. This study developed and validated a test that could be used to test students level of integrated science process skills of i. formulating hypotheses, ii. defining variable operationally, iii. identifying and controlling variables, iv. planning investigations as well as v. analyzing and interpreting data. This provides an opportunity for another study to develop and validate instruments which will test students' competence in basic science process skills basic science process skills such as observing, asking questions, classifying, measuring, and predicting based on the format, procedures, and methodology used in this study.

iii. The use of a paper and pencil test to assess science process skills like the developed BPST has been criticized by several researchers who advocate for practical manipulation of apparatus and physical demonstration of practical skills. This weakness is ana opportunity for other researchers to develop the authentic and practical models that will provide students greater opportunities for students to demonstrate competence in science process skills, such as lab-based assessments or and exploration-based tests.

8.3 Second phase: assessing the knowledge level of science process skills of Biology students in the municipality of Morogoro

8.3.1 Summary of findings and conclusion

It is eleven years now since the inception of the competence-based curriculum in Tanzania. The newly revised Competence Based Curriculum of 2005 had placed a heavy emphasis on the need for secondary school science learners to acquire science process skills. However, there was no clear evidence of whether or not learners are appropriately acquiring competence in these scientific skills as prescribed in the curriculum. The phase of the study was conducted in order to establish a base level of information Tanzania student´s knowledge level of the science process skills. The mini study employed the test (BPST) that was been validated test in the first stage to examine the knowledge level of science process skills of advanced level secondary school Biology students in Morogoro municipality. The analysis was also done to examine the performance of students in individual integrated science process skills (formulating hypotheses, defining variable operationally, identifying and controlling variables, planning investigations and analyzing and interpreting data) and determining whether there was a statistical difference in performance of students based on individual science process skill.

Based on the Biology process skills test (BPST) scores, it was found that Biology students in Morogoro municipality had barely average knowledge level of integrated science process skills. The mean of test scores was 17.2 items out of 35 items in the test corresponding to 49.1%. However, Morogoro students performed relative better on items measuring their ability in identifying and controlling variables with score mean of 4.05 out of 07 items and they performed extremely poor on items which measured their skills in analyzing and interpreting data with the mean of 2.34 out of 07 items. Due to the influence of social forces, culture and gender roles in the Tanzania, anecdotal evidence would suggest male students have higher levels of achievement in science-related disciplines than females. However, the findings from Morogoro Biology students in this study did not support that assertion. Based on the science process skills test scores of the 246 females and 107 males in the study, the female had a relatively higher mean of 17.75 than their counterpart male who had a score mean of 15.76. Statistical significant differences were found to exist between male and female students in terms of their performance in science process skills through SPSS t-test. Although experience and maturity might be strong determinants for one's academic performance, an independent t-test of students score means based on their years of schooling in this study failed to find a statistically significant difference of Form V and Form VI students. Form VI has one extra year in the education system than their counterpart Form V. The mean of 142 Form VI who participated in this study was 17.37 while the mean of 211 Form V students was 17.00.

8.3.2 Specific Recommendations for Biology teachers and science education researchers in Tanzania as a result of findings in this stage

i. Teaching science process skills is a challenging experience but a rewarding one. Science teachers need to capitalize on opportunities for classroom activities. While this is not an easy solution, it remains the best alternative available at this time where there is a serious shortage of textbooks emphasizing and guiding the teaching of science process skills. Teachers should give as many authentic assignments to students as possible. This will help students to develop scientific thinking and skills.

ii. The finding discussed in this study on the level of science process skills of Morogoro students has been have collected from the developed multiple choices Biology process skills test (BPST). However, students' competency in integrated science process needs to be measured directly and authentically. That means for realistic assessment students are required to demonstrate in hands-on experiment or investigation. The researcher recommends other forms of assessment open-ended questions and performance tests or to supplement the findings gathered from the multiple choice test. Although it would be very time-consuming to directly measure all these skills in science classroom activities. However, the written test may not tell if learners would be able to display the process skill when required.

iii. Knowledge level of integrated science process skills among the advanced level Biology students in Morogoro was found to be unsatisfactory. It is a pedagogical triumph to teach scientific fact in relation to the integrated process skill. The researcher recommends that Tanzania science teachers must understand how to guide pupils to use these processes in the context of learning science concepts. In the teaching of science process skills, science teacher's main job is to provide the situation for the hands-on learning rather than to tell about or explain theses skills. Research studies that deal with student ability to investigate the call for educational practitioners to pay attention to incorporating the skills of scientific inquiry in the process of teaching and learning.

iv. The overall level of performance of Morogoro students in this study cannot be regarded as good considering the high premium placed on the students' acquisition of science process skills in their new science curricula. If these skills are not developed sufficiently, students cannot be good future scientists and researchers. For this reason, the basic target in science classes should be teaching students how to attain knowledge rather than passing the convenient knowledge. In so doing teachers should not expect mastery of experimenting skills after only a few practice sessions. Instead, students need multiple opportunities to work with these skills in different content areas and contexts.

8.3.3 Recommendations to the Ministry of Education and Vocational Training officials of Tanzania

The following are some recommendations to the ministry of education and vocational training which is responsible for the formulation, monitoring, and evaluation of the implementation education policies in Tanzania.

i. The result on the knowledge level of Morogoro Biology students from this study showed that low level of mastery of the science process skills by students. The study recommends that, in line with the emphasis on the teaching of science process skills for the secondary school science curriculum, the Ministry of Education and Vocation Training in Tanzania through its National Examination Council of Tanzania (NECTA) needs also to revamp the standardized tests. The assessment of integrated process skill in a written format should be introduced in the public examination besides the school based laboratory assessment. Items in the written test required students to plan and design an investigation that involves the use of all the integrated process skills. To answer items correctly, students must possess all components of integrated process skill as well as knowledge in science. Students who are weak in the content area may not be able to apply these skills.

ii. Unsatisfactory performance of Morogoro students in science process skills test might also attribute the weakness of their teachers in this area. (i) Pre-service science teacher preparation programs are in a unique position to enhance the teaching capacity of prospective science teachers. Hence the study recommends to the ministry that, effective science instruction methods in teacher training institutions of all levels should be infused into teaching methods courses as a way to reinforce the scientific rigor to prospective science teachers. Moreover, method-based and hands-on investigative activities should be a significant component of pre-service teacher education. This will help future science teachers to realize that they have the potential capabilities, ability, and capacity to teach their students effectively in the area of science process skills.

iii. It is further recommended that the information obtained from this study should be useful in planning appropriate teacher training program so as to prepare science teachers in the implementation of the science education in school.

iv. It is also recommended that the Ministry of Education and Vocational Training of Tanzania should apply the findings from this study to encourage discussions among science education actors about desired student outcomes in science and assessments appropriate to those outcomes and the appropriate techniques to the teaching of scientific skills.

8.3.4 Limitations of the study in this stage (second stage)

Although significant findings have been presented and discussed with respect to the knowledge level of Tanzania students in the areas of science process skills, the limitations regarding the practical significance of these findings need to be discussed.

i. This study was confined to only 4 secondary schools in Morogoro municipality using 353 Biology students. However, other researchers may work and involve bigger samples countrywide in order to understand the phenomena under study and to attempt generalization through replication of findings.

ii. For students to demonstrate the integrated process skills, assessment using hands-on procedures (authentic assessment) to determine skill acquisition by students is pivotal. The use of a paper and pencil test to assess practical skills like BPST has been criticized by several researchers who advocate for practical manipulation of apparatus and physical demonstration of practical skills. This presents a limitation in the findings of this study in the sense that the instrument used in this study (BPST) does to accommodate these requirements.

8.3.5 Areas recommended for further research

Generalization of these findings is not possible because of the nature of the study and the sample size used. However, the findings from this study on the knowledge level of scientific skills of Morogoro students present several research opportunities in the following areas;

i. This study was confined to only 4 secondary schools in Morogoro municipality using 353 Biology students. However, other researchers may work and involve bigger samples countrywide in order to understand the phenomena under study and to attempt generalization through replication of findings.

ii. Science process skills present a new research topic in Tanzania education system. This current study focused on assessing Biology students' knowledge and competence in the area of integrated science process skills in Morogoro. However, several studies should also be conducted to assess students' knowledge level of science process skills at different education levels (primary, secondary, and tertiary). These studies will provide directions as to whether pupils and students at all education levels are acquiring science process skills as campaigned by the 2005 Competence Based Curriculum.

iii. The findings that Morogoro students have unsatisfactory knowledge level of science process skills implies that the teaching of content must take precedent over the training of students on the acquisition of science process skills. This gives an opportunity for further studies especially in investigating what actually happens in the science classrooms to shed some light on the effectiveness of the teaching of the review science curriculum at the secondary school level.

iv. This study on the assessing the knowledge level of science process skills of advanced level Biology students in the municipality of Morogoro opens the opportunity for another study to determine Tanzania teachers' competence in integrated science process skills. The information about teachers' knowledge level of science process skills is critical because the competence of students in this area depends on the competence of their teachers. No education system can rise above the quality of its teachers. Hence teachers' satisfactory capacity in science process skills is a key to successful implementation of the new science curriculum which emphasizes these scientific skills.

8.4 Third phase: The effectiveness of inquiry based approach in the development of students' science process skills, conceptual understanding and motivation

8.4.1 Summary of findings and conclusion

This part of research compared the effectiveness of the inquiry-based approach and conventional method in the development of students' science process skills, conceptual understanding of genetics and in enhancing students' motivation towards science process skills. Motivational constructs under study included i. grade motivation, ii. intrinsic motivation, iii. self-efficacy, iv. self-determination, v. career motivation, and v. self-concept towards science. The study also correlated students' achievement in science process skills with their development in i. conceptual understanding, ii. intrinsic motivation, iii. self-efficacy, iv. self-determination, and v. self-concept.

The analysis of post process skills revealed that the experimental group students performed better after undergoing the experimental treatment of inquiry constructivist approach as compared to the scores of the control group. An analysis of independent samples t-test based on the type of instruction students received at (α) =0.05 produced a p of 0.047 and a t value of 0.633, hence rejecting the null hypothesis (Ho1). An analysis of student gains on the genetic test between a group of traditionally-taught students and a group of students taught by inquiry-based methods was conducted to show if there are positive effects of inquiry in the realm of Biology. The findings suggest that there is no statistical difference in students' understanding of themes within genetics after being exposed to different methods of instruction. This is contrary from the way it was hypothesized, inquiry-based instruction, compared with expository instruction, did not cause a statistically significantly better acquisition of scientific concepts related to genetics. Although students in both groups learned adequate genetics, but there was no statistically significant difference in the performance of students in the experimental group and the control group. Possible causes for this deviation from the previously cited literature could have arisen as a consequence of varying factors such as; the quantity of information to be disseminated, students' familiarity with the lecturer (a researcher), students previous exposure to both teaching methods, students' preparedness to learn in a an inquiry classroom, students' educational cultural context, students' past experience with the rudiments of motivation, and the nature of the subject content.

Findings show that both groups were equal in their initial status (before intervention) on six motivation constructs which included i. intrinsic motivation, ii. grade motivation, iii. career motivation, iv. self-efficacy, v. self-determination, and vi. self-concept based on the t-test and descriptive statistics. Contrary from the researcher's conception, no evidence that the experimental group taught genetics through inquiry approach achieved more on the posttest on motivation as compared with a control group from descriptive and inferential statistics of with t- test. It was expected that the two groups of students would differ in terms of their motivation towards science process skills at posttest in favor of the experimental group. This is because, throughout intervention using genetics lessons, these students in the experimental group were working in small groups where they were encouraged to explore problems, formulate hypotheses, designing micro experiments share their ideas with their classmates, discuss their observations and interpret findings of their hands-on activity carried out.

It is concluded that these findings are preliminary and should be treated with caution. At first glance, the results seem to invoke notions of the Hawthorne effect, an experimental effect in the direction expected but not for the reason expected; i.e. a significant positive effect that turns out to have no causal basis in the theoretical motivation for the intervention, but is apparently due to the effect on the participants of knowing themselves to be studied in connection with the outcomes measured. However, implications for successful implementation in areas of initial delivery, remediation and enrichment are bountiful and imply that student interest and motivation towards science process skills can be increased with the inquiry activities.

8.4.2 Based on the findings, the following recommendations are made for Tannzania science teachers:

i. With science process skills performance, a statistically significant difference in posttest scores between students in the experimental group in which inquiry-based instruction approach was applied and the control group taught traditionally in the BPST in favor of the experimental group was seen. This means that scientific process skills can be improved in time with the help of student activities as in inquiry-based approach. Tanzania science teachers need to teach science students in the same way science is done by scientists if they have to acquire science process skills (ie by explore problems, formulate hypotheses, designing micro experiments share their ideas with their classmates, discuss their observations and interpret findings).

ii. It should be noted that the inquiry-based approach to science especially experiment designing takes more time and energy both to the part of students and teachers. In the Tanzania, a typical teacher teaches from four to six hours per day. Considering this, teachers have limited time to prepare inquiry-based lessons.Hence the study recommends teachers to use time efficiently and at an optimal level so that they have enough time for students to participate in experimental activities in practical applications. On the other hand improved theoretical models via discussion-centered experimental activity papers are also recommended.

iii. It was concluded in the study that the application of inquiry-based instruction approach had positive effects on the scientific process skills hence Tanzania teachers should encourage students can to access true knowledge by allowing them to research and question instead of presenting them readily available and acceptable knowledge. The development of skills in scientific inquiry requires that students of science be provided with appropriate and adequate guidance in their study of science. Inquiry-based instruction includes a variety of teaching strategies, such as questioning; focusing on language; and guiding students to make comparisons, analyze, synthesize, and model

iv. Findings on genetics performance suggest that students in both groups learned adequate genetics and that there was no statistical difference in students' understanding of genetics after being exposed to different methods of instruction at posttest. However, for long-term retention of knowledge, Biology teachers are recommended to use student-centered inquiry approach in which there is an active student involvement in the lessons, answering questions, and interacting with each other and with their teachers. Therefore teachers should provide a learning environment that fosters scientific reasoning and encourages students to develop a meaningful learning approach.

v. Tanzania science teachers need to raise students' intrinsic motivation, personal relevance, self-determination, and self-efficacy to enhance better learning outcomes in combined science. Perhaps, the most important of all, teachers should teach combined science in such a way that it is interesting and enjoyable for students. Tanzania teachers should create a conducive learning environment that is challenging, stimulating and relevant to boost students' interest and motivation, for instance, promoting cohesiveness among students using small group cooperative learning strategies. This is a powerful pedagogical tool that enhances students' motivation towards science and science process skills in specific.

vi. Teachers should explore and use this strategy to make students more determined and efficacious to learn combined science instead of using the teacher-centered expository approach that is so prevalent among science teachers. Teachers should also attempt to link science concepts to students' experiences, so that they can realize the relevance of what they learn to their everyday lives, thus making learning more meaningful and relevant. When teachers are able to foster and increase students' motivation to learn the processes of science, it is likely that many more students will be able to successfully complete their secondary education with scientific competence.

8.4.3 Recommendations for the Ministry of Education and Vocational Training with respect to teaching through inquiry-based approach

i. Although in this study the inquiry-based approach had the same superiority in enhancing motivation and knowledge of Morogoro students, but its application was found to have positive effects on the development of scientific process skills. Hence it is recommended that effective science instruction methods in teacher training institutions of all levels should be infused into teaching methods courses as a way to reinforce the scientific rigor of prospective science teachers. The inquiry method-based and hands-on investigative activities should be a significant component of pre-service teacher education. This will help future science teachers to easily teach their students using this approach which has also been suggested by the new curriculum.

ii. The Ministry of Education and Vocational Training should play a critical should help school teachers to enhance inquiry teaching by supporting secondary schools to successfully integrate laboratory and inquiry-based experiences with science instruction and enhance science process skills acquisition to students. This can be done by providing schools with adequate and improved laboratory equipment, reagents, apparatuses and supporting staff like laboratory technicians.

iii. The Tanzania Institute of Education (TIE) as a curriculum development unit should (a) produce new science textbooks in conjunction with new syllabuses and science process skills, and (b) develop a set of laboratory activities to guide secondary school science teachers in teaching science and science process skills using inquiry approach. The need for books which depict the Tanzania environment (e.g., biological sciences) is very essential. The staff of the institute of curriculum development in collaboration with the subject teachers in various schools with assistance from university instructors should prepare pamphlets or books for use in schools. University instructors should publish subject books in disciplines.

iv. Improving science teachers' competence in laboratory experiences (activities) is of critical importance if integrated science process skills are to be well taught. Teacher training institutions which are charged with the responsibility of training science teachers at all education levels in Tanzania are increasingly producing teachers without requisite laboratory practical experiences. Thus, ongoing science teacher education and professional development programs should ensure that prospective science teachers have practical experiences that familiarize them with inquiry-based laboratory skills needed to facilitate the teaching of science process skills. This requires major and fundamental changes in pre-service science teacher education contents, including providing a range of effective laboratory experiences.

v. Teachers orientation on the requirements of the competence-based curriculum, science process skills and hence inquiry-based approach to science should be a matter of urgency. It was observed from this study that, many teachers have not yet received any kind of training on the competency-based approaches such as inquiry. Hence, the Ministry of Education and Vocational Training should organize and coordinate demand driven in-service training for teachers about the requirements of the new curriculum.

vi. Science teacher education programs in all Tanzanian tertiary institutions should be accredited to ensure that teacher educators are competent and knowledgeable and that all pre-service teachers have sufficient teaching practice and graduates meet necessary standards for entry to the profession. Graduating teachers should be equipped with the various strategies for conducting inquiry-based teaching and formative assessment practices during their initial teacher education training so as to improve teaching and learning in science.

vii. There should be a focus on opportunities for ongoing professional learning for all categories of teachers including in-service science teachers. The Ministry of Education and Vocational Training as a standard regulatory unit should mount workshops for in-service science teachers to update them on the teaching through an inquiry-based approach. In addition, a system of ongoing, school-based in-service teacher training is needed to allow teachers to continually develop and upgrade their pedagogical skills and content knowledge.

viii. The ministry of education Tanzania should make sure schools are equipped with appropriate teaching and learning materials. It has been observed that the lack of appropriate teaching materials is another prevalent problem for the implementation of inquiry-based teaching. Most teaching materials in the Tanzania are written in an expository way. As such, students often view science as accumulated facts rather than a form of investigation. With these kinds of resources, students find difficulties translating the science ideas contained in their textbook into an active form of inquiry. In the Tanzania, most classrooms and schools are not equipped with sufficient teaching materials and do not serve as venues to engage students in hands-on or minds-on types of inquiry learning. School climate, teachers' expertise, and availability of inquiry-based materials are therefore critical in the effective implementation of inquiry-based teaching.

8.4.4 Limitations of this quasi-experimental study

These findings from the quasi-experimental study to assess the effectiveness of the inquiry-based approach in the development of students' science process skills, conceptual understanding of contents and in enhancing motivation are preliminary and should be treated with caution. The study does not escape the limitations of similar quasi-experimental studies in science education research. Generalization of these findings is not possible because of the nature of the study and other methodological weaknesses as follows,

i. The study with the experimental group using inquiry-based approach was carried out in relatively crowded classrooms. The average population of the students per classroom was about 50 students. Bigger classes reduce the number of time students can actively engage with each other, increase the disruptive behavior in the classroom and reduce the amount of time the teacher can spend working with each individual student. Hands-on activities and lessons were given under crowded classes. In this respect, therefore, any generalizations from these findings need to be taken cautiously.

ii. The second limiting factor was the lack of enough participants to make a control group. This study was confined to only 3 secondary schools in Morogoro municipality using 263 Biology students of which of which 169 constituted the experimental group while only 69 constituted the control group. Generalization of findings would only be possible if the samples size was bigger and countrywide in order to understand the phenomena under study.

iii. It should be noted that the experiment designing process skill takes more time and energy to improve. In this study, there was not enough time for students to participate in experimental activities in practical applications. Hence, some experimental activities that could be performed in a laboratory to determine the effects of this teaching approach on experiment design were conducted theoretically via discussion-centered experimental activity papers.

iv. Another weakness of this study is the fact that it didn't analyze the performance of participating students based on their sex although the data about their sex were collected. With the inclusion of this statistical information, many of this study's findings could have lead to further discussion as to the impact this intervention could have had on boys versus girls.

v. In addition, because all students in this study were boarding school students (living and sleeping in their respective schools), it is possible that students in both groups communicated with each other regarding the lessons and tests were given. It is impossible and unethical to isolate all of the participants completely. Each school in the study had both groups (ie Form six as a control group and Form five as an experimental group) of hence it is reasonable to assume that they mix outside of lessons and share ideas, potentially contaminating the results. Although this might influence the results by reducing intervention effects and therefore providing further support for any significance reported in this study, it is possible that communication between students heightened the both students' awareness of the intervention, thus positively affecting their science process skills, genetics knowledge and related motivations.

vi. First, it is possible that group differences in the science process skills, genetics and or genetics may have resulted from teacher effects and other confounding background factors not explored in this study. Although efforts were made to reduce confounding variables, the authors do not ignore the possibility of differences in treatment compliance regarding the guidelines for implementing both the inquiry lessons and control group lessons.

vii. Another weakness of this study is the fact that no follow-up study was conducted to compare retention of students in both groups with respect their knowledge level of science process skills, conceptual understanding of genetics and in motivation. Follow-up study would enable the present researcher to find out whether there has been a change in the variables studied since the initial study.

8.4.5 Recommendations on Areas for Further Research

Overcoming the potential limitations of this study provides guidance for further research. However, the findings from this study present several research opportunities. These research opportunities are in the following areas;

i. Firstly, this study was based on a sample from three (03) secondary schools in Morogoro with 263 students only. This suggests that replication in alternative educational settings is needed for greater generalization. Another research works involving bigger samples countrywide are needed to understand the phenomena under study and to attempt generalization through replication of findings. It is possible that the sample involved in this study was not large enough to detect significant trends. This is an excellent opportunity for further research.

ii. Second, longer instructional periods may be needed for this kind of quasi-experimental study in order to accomplish the development of scientific skills, conceptual understanding, and motivational beliefs. This study took only two months for each school (a total of six months in three schools involved). Longitudinal studies may be essential in this respect. Hence this study opens an opportunity for a similar study but to be conducted as a longitudinal research. It is possible that the results may vary with the longevity of the application period of the treatment (inquiry approach and traditional methods).

iii. Science process skills seem a new research topic in Tanzania education system. This current study focused on assessing Biology teachers' knowledge and competence in the area of integrated science process skills. However, several studies should also be conducted to assess the effectiveness of the inquiry-based approach in the acquisition of science process skills at different education levels (primary, secondary, and tertiary). These studies will provide directions as to whether pupils and students at all education levels are acquiring science process skills as a result of applying the inquiry approach.

iv. Fourth, this study focuses on guided inquiry, in which students were given lesson sheets and required to identify the scientific problem, analyze data, formulate hypotheses, and explain the Biological phenomenon for the basis of the lesson. They were given opportunities to engage, explore, explain, extend (or elaborate), and evaluate the lessons. However, this study on guided inquiry model opens an opportunity for a similar study but to be conducted as an open inquiry activity. Open inquiry is the highest level of inquiry in which students in their own simulate and demonstrates ownership and responsibility of the lessons and it requires high order thinking capabilities than the guided one.

v. Another research opportunity would be to expand the study to other grade levels and into other subjects apart from Biology (genetics). Results may vary based on the grade level, demographics, or subject. By expanding the study, further research could be conducted examining if a correlation exists between the age of the students and their development of science process skills, conceptual understanding, and motivation.

vi. Another opportunity as a future research would be comparing student self-report measures such as tests and Likert scales which was used in this research to study motivation constructs with the qualitative results from the interviews and classroom observations. In using self-report measures (like SMQ-II and FSWEx scales), this research assumes that students have the ability to verbally express their motivations and cognitions. However, it is possible that students are incapable of identifying and recalling their mental processes. Students may be using motivated in one of the constructs but not possessing the conscious awareness that they are doing so. Hence this study is an opportunity for another motivation study that will integrate both student self-report measures such as tests and Likert scales with the qualitative results from the interviews and classroom observations.

vii. On motivation, this study assessed the effectiveness of inquiry-based approach on six motivational constructs under as defined by the proponents of social-cognitive theory which included i. grade motivation, ii. intrinsic motivation, iii. self-efficacy, iv. self-determination, v. career motivation, and v. self-concept. However other researchers could also consider employing structural equation modeling to determine the relationships between students' motivational constructs and their performance in combined science. Structural equation modeling (SEM) refers to a diverse set of mathematical models, computer algorithms, and statistical methods that fit networks of constructs to data.This would provide vital information on the variance of the different motivational constructs on students' achievement which may be influenced by factors, such as grade level, gender, ethnicity, and subject area.

viii. Due to the scope of this study and its interdisciplinary nature, it would best be conducted as a research funded project involving longer time and several teams such as statisticians, researcher, Biologists and educationists. Interdisciplinary research inquiries critically draw upon two or more disciplines and which lead to an integration of disciplinary insight. This requires effort and investment of time. Hence the researcher recommends the Ministry of Education and Vocational Training in Tanzania to find the means to fund the research of this nature which will be conducted as an interdisciplinary research funded project in order to understand the phenomena under study from different perspectives.

REFERENCES

Abu-Hilal, M. M., & Bahri, T. M. (2000). Self-concept: The generalizability of research on the SDQ, Marsh/Shavelson model and I/E frame of reference model to United Arab Emirates students. *Social Behavior and Personality*, 28, 309-322.

Adey, P., & Shayer, M. (1990). Accelerating the development of formal thinking in middle and high school students. *Journal of Research in Science Teaching,* 27(3), 267–285.

Adkins, D. (1974). *Test construction: Development and interpretation of achievement tests*. Columbus, Ohio: Charles E. Merrill Publishing Company

Afuwape. A. (2011). Students' self-concept and their achievement in basic science. *Journal of African Research Review;* 5 (4), 191-200

Alberts, B. (2000). *Some thoughts of a scientist on inquiry.* In J. Minstrell, & E. H. Zeevan (Eds.), Inquiring into Inquiry Learning and Teaching in Science. Washington DC: American Association for the Advancement of Science.

Allen, M., & Yen, W. (2001). *Introduction to measurement theory*. New York: Waveland Press.

Alkin, M. (1969). *Mathematics 4-6, instructional objectives exchange*. California Univiversity., Los Angeles. Center for the Study of Evaluation. Office of Education (DHFW),Washington, D.C. Bureauof Research. Retrieved from http://files.eric.ed.gov/fulltext/ED034702.pdf

Akcay, H & R. Yager, R. (2016). Students learning to use the skills used by practicing scientists. *Eurasia Journal of Mathematics, Science & Technology Education,* 12(3), 513-525

American Association for the Advancement of Science (AAAS) (1993). *Benchmark for Science Literacy: Project 2061*. New York: Oxford University Press. Retrieved from http://www.project2061.org/tools/benchol/bolframe.html

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Anderson, L.W., Krathwohl, D.R., Airasian, P.W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., Raths, J., Wittrock, M.C. (2001). *A taxonomy for learning, teaching, and assessing: A revision of bloom's taxonomy of educational objectives*. New York: Pearson, Allyn & Bacon.

Anderson, R.D. (2002).Reforming science teaching: What research says about inquiry?. *Journal of Science Teacher Education*, 13(1), 1-12.

Angleitner, A., John, O.P., and Löhr, F.J. (1986). It's what you ask and how you ask it: An itemmetric analysis of personality questionnaires. In A. Angleitner and J. S.

Wiggins (Eds.), Personality assessment via questionnaires. Current issues in theory and measurement (pp. 61-108). Berlin, Germany: Springer. Downloaded from https://pub.uni-bielefeld.de/publication/1779424

Anthoine, E; Moret, L; Regnault, A; Sobille, V and Hardouin, J. (2014). Sample size used to validate a scale: a review of publications on newly-developed patient reported outcomes measures, *Journal of Health and Quality of Life Outcomes*,12:176 http://www.hqlo.com/content/12/1/176

Arena, P. (1996). The role of relevance in the acquisition of science process skills. *Australian Science Teachers Journal*, 42 (4), 24-38.

Ary, D., Jacobs, L., & Razavieh, A. (1990). *Introduction to research in education*. Fort Worth: Harcourt Brace College.

Ashton, P., & Webb, R. (1984). Teacher efficacy: A motivational paradigm for effective teacher education. *Journal of Teacher Education*, 35 (5), 28-32.

Ashton, P., Webb, R., & Doda, C. (1983). *A study of teachers' sense of self-efficacy*. Gainesville: University of Florida Press; downloaded at 3rd August 2014 from http://files.eric.ed.gov/fulltext/ED231833.pdf

Ateş, S. (2004). The effects of inquiry-based instruction on the development of integrated science process skills in trainee primary School teachers with different piagetian developmental levels. *Journal of Research in Science Teaching* 24, (3) 275-290.

Athuman, J. (2010). *An assessment of science teachers' capacity for teaching integrated science process skills: A case of secondary school biology teachers in Morogoro municipality*. Unpublished Master Dissertation. University of Dar es salaam

Ayala, C.A., Shavelson, R.J., Yin, Y., & Schultz, S. (2002). Reasoning dimensions underlying science achievement: The case of performance assessment. *Journal of Educational Assessment*, 8(2), 101-122.

Bachman, L. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Journal of Language Testing*, (17), pp. 1–42.

Bailin, A., & Grafstein, A. (2001). The linguistic assumptions underlying readability formulae: A critique. *Journal of Language and Communication*, 21, 285-301.

Baird, W.E. and Borich, G.D. (1987). Validity consideration for research on integrated science process skills and formal reasoning ability. *Journal of Science Education,* 71(2), 259-269

Baker, D., & Leary, R. (1995). Letting girls speak out about science*. Journal of Research in Science Teaching*, 32(1), 3-27.

Balci, S., Cakiroglu, J., & Tekkaya, C. (2006). Engagement, exploration, explanation, extension, and evaluation (5E) learning cycle and conceptual change text as learning tools. *Journal of Biochemistry and Molecular Biology Education*, 34, 199-203.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory* Prentice-Hall, Inc, Englewood Cliffs, NJ. Retrieved from http://search.proquest.com.ezaccess.libraries.psu.edu/docview/617099314?accountid=13158

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology,* 52(1), 1-26 . Retrieved from http://www.annualreviews.org/doi/abs/10.1146/annurev.psych.52.1.1

Bandura, A. (1997). Self-efficacy: *The exercise of control*. New York: Freeman

Banet, E., & Ayuso, E. (2000). Teaching genetics at secondary school: A strategy for teaching about the location of inheritance information. *Journal of Science Education,* 84, 313-351.

Barakatas, A. (2005). A typology of mathematics teachers' beliefs about teaching and learning mathematics and instructional practices. *Mathematics Educational Research Journal.* 17(2), 69-90.

Barman, C. R., Barman, N. S., & Miller, J. A. (1996). Two teaching methods and students' understanding of sound. *Journal of School Science and Mathematics*, 96, 63-67.

Barmby, P., Kind, P. M., & Jones, K. (2008). Examining Changing Attitudes in Secondary School Science. *International Journal of Science Education*, 30(8), 1075 - 1093.

Barnett, D. W., Daly, E. J., Jones, K. M., & Lentz, E. F. (2004). Response to intervention: Empirically based special service decisions for single case designs of increase and decreasing intensity. *Journal of Special Education* , 38(1) , 66–81.

Bauer, Henry H. (1992). *Scientific literacy and the myth of the scientific method.* Urbana and Chicago: University of Illinois Press.

Bazler, J. A., & Simonis, D. A. (1991). Are high school chemistry textbooks gender fair?. *Journal of Research in Science Teaching*, 28(4), 353-362.

Beaumont-Walters, Y., & Soyibo, K. (2001). An analysis of high school students' performance on five integrated Science process skills. *Research in Science & Technological Education*, 19(2), 133-145.

Beisenherz, P. and Dantonio, M. (1996). *Using the learning cycle to teach physical science: A hands-on approach for middle grades*. Portsmouth, NH Heinemann

Bellmore, A.D., & Cillessen, A.R.N. (2006). Reciprocal influences of victimization, perceived social preferences, and self-concept in adolescence. *Journal of Self and Identity*, 5, 209-229.

Benabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *Journal of Review of Economic Studies*, 70, 489-520

Ben-Zur, H., & Zeidner, M. (1989). Sex differences in anxiety, curiosity and anger: A cross-cultural study. *Journal of Sex Roles*, 19, 335–347.

Berg, C. A. R., Bergendahl, V. C. B., & Lundberg, B. K. S. (2003). Benefiting from an open-ended experiment? A comparison of attitudes to, and outcomes of, an expository versus an open-inquiry version of the same experiment. *International Journal of Science Education,* 25(3), 351-372.

Berg, C. A., & Philips, D. G. (1994). An investigation of the relationship between logical thinking structures and the ability to construct and interpret line graphs. *Journal of Research in Science Teaching*, 31, 323–344.

Berliner, D. C. (1986). In pursuit of the expert pedagogue. *Journal of Educational Researcher, 15*(7), 5–13.

Bilgin, I. (2006). The effects of hands-on activities incorporating a cooperative learning approach on eight grade students' science process skills and attitudes toward science. *Journal of Baltic Science Education,* 1(9). 27-37.

Bloom, B.S. (Ed.). Engelhart, M.D., Furst, E.J., Hill, W.H., Krathwohl, D.R. (1956). *Taxonomy of educational objectives, handbook I: The cognitive domain*. New York: David McKay Co Inc:
Retrieved from https://de.scribd.com/doc/53565531/Taxonomy-of-Educational-Objectives-Handbook-1-Cognitive-Domain-0582280109

Blosser, P. (1994). What research says about achievement in science?. *Journal of School Science and Mathematics*, 84, 514-521.

Bolliger, D., Supanakorn, S., Boggs, C., (2010).Impact of podcasting on student motivation in the online learning environment. *Journal of Computers and Education*, 55 (2), 714-722.

Bond, L. (1996). Norm- and criterion-referenced testing. *Practical Assessment, Research & Evaluation, 5*(2). Retrieved from http://ericae.net/pare/getvn.asp?v=5&n=2.

Boone, H. N., Jr. (1990). Effect of level of problem solving approach to teaching on student achievement and retention. Journal of Agricultural Education, 31(1), 18–26. doi:10.5032/jae.1990.01018

Borg, W., & Gall, M. (1989). *Educational research: An introduction*. New York: Longman.

Boudreaux, A., Shaffer, P., Heron, P., & McDermott, L. (2008). Student understanding of control of variables: Deciding whether or not a variable influences the behavior of a system. *American Journal of Physics*, 76, 163-170. http://dx.doi.org/10.1119/1.2805235

Bowling, B. V. (2007). Development, evaluation, and use of a genetic literacy concept inventory for undergraduates (Doctoral dissertation). University of Cincinnati, USA: Retrived from https://etd.ohiolink.edu/rws_etd/document/get/ucin1185456791/inline

Bowling, B.V., Acra, E.E., Wang, L., Myers, M.F., Dean, G.A., Markle, G.C., Moskalik, C.L., & Huether, C.A. (2008). Development and evaluation of a genetics literacy assessment instrument for undergraduates. *Journal of Genetics,* 178, 15-22. doi:10.1534/genetics.107.079533.

Bradley, J. D. (2000). *The Microscience project and its impact on pre-service and in-service teacher education.* Washington, DC: The World Bank.

Brant, R. (2000). On performance assessment: A conversation with Grant Wiggins. *Journal of Educational Leadership, 49*(*8*), *35-37*.

Braun, H. I. (1988) Understanding score reliability: Experiments in calibrating essay readers. *Journal of Educational Statistics*, 13 (1), 1-18.

Bredderman, T. (1983). Effects of activity based elementary science on student outcomes: A qualitative synthesis. *Journal of the Review of Educational Research*, 53 (4), 499-518.

Brendgen, L .(2002). Overcoming low self-esteem: A cognitive behavioral approach*. Journal of Stress News*, 14, 7-11.

Brew, A. (2003). Teaching and research: New relationships and their implications for inquiry-based teaching and learning in higher education. *Journal of Higher Education Research and Development* 22(1): 3-18.

Brooks, J.G., & Brooks, M.G. (1999). In search of understanding: The case for constructivist classrooms. Alexandria, VA: Association for Supervision and Curriculum Development.

Brophy, J., & Evertson, C. (1981). *Student characteristics and teaching*. New York: Longman.

Brotheton, P. N., & Preece, P.F.(1995). Science process skills: Their nature and interrelationships. *Journal of Research in Science & Technological Education*, 13 (1), 5-12.

Broussard, S. C., & Garrison, M. E. (2004). The relationship between classroom motivation and academic achievement in elementary school-aged children. *Family Consumer Science Research Journal*, 33(2), 106-120.

Brown, J. D. (2000). Statistics corner. Questions and answers about language testing statistics (How can we calculate item statistics for weighted items?). *Journal of Testing & Evaluation,* 3 (2), 19-21. Retrieved from http://jalt.org/test/bro_6.htm

Bülent, A. (2015). Examining preservice science teachers' skills of formulating hypotheses and identifying variables. *Journal of Asia-Pacific Forum on Science Learning and Teaching,* 16 (1):
Downloaded from http://eric.ed.gov/?id=EJ1070729

Burns, J. C., Okey, J. R., & Wise K. C. (1985). Development of an integrated process skills test: TIPS II. Journal of Research in Science Teaching, 22 (2), 169-177.

Burris, S. & Garton, B. L. (2007). Effect of instructional strategy on critical thinking and content knowledge: Using problem-based learning in the secondary classroom. *Journal of Agricultural Education*. 48(1), 106-115

Burton, L & Mazerolle, S. (2011). Survey instrument validity part I: principles of survey instrument development and validation in athletic training education research. *Athletic Training Education Journal*, 6(1):27-35.

Bybee, R. W., Taylor, J. A., Gardner, A., Van Scotter, P., Carlson Powell, J., Westbrook, A., & Landes, N. (2006). *The BSCS 5E Instructional model: Origins, effectiveness and applications.* Retrieved from **http://www.bscs.org/bscs-5e-instructional-model**

Bybee, R. W. (2012). The next generation of science standards: Implications for biologyeducation*. The American Biology Teacher,* 74 (8), 542–549.

Byrne, B. (1990). Self-concept and academic achievement: Investigating their importance as discriminators of academic track membership in high school. *Canadian Journal of Education*, 15(2), 173-182.

Campbell, R & Martinez-Perez, L. (1976). A Study of Relationships of science attitudes, achievement and self-concept of pre-Service teachers. Downloaded from http://files.eric.ed.gov/fulltext/ED125898.pdf

Cardoso, D. C., Cristiano, M. P. & Arent, C. O. (2009). Development of new didactic materials for teachingvscience and biology: The importance of new education practices. *OnLine Journal of Biological Sciences*, 9(1), 1-5.

Carin, A. A., & Bass, J. E. (2001). *Teaching science as inquiry*. Upper Saddle River, NJ: Prentice-Hall.

Cassady, J.C., & Johnson, R.E. (2002). Cognitive test anxiety and academic performance. Journal of Contemporary Educational Psychology, 27, 270–295.

Cavallo, A. M. L. (1996). Meaningful learning, reasoning ability, and students' understanding and problem solving of topics in genetics. *Journal of Research in Science Teaching*, 33(1), 625-656.

Chall, J. S. & Dale, E. (1995). *Readability revisited: The new Dale-Chall readability formula*. Cambridge, MA: Brookline Books. Retrived from http://www.readabilityformulas.com/new-dale-chall-readability-formula.php

Chall, J. S. (2000). *The academic achievement challenge.* New York: Guilford.

Change, C., & Mao, S. (1998). *The effects of an inquiry-based instructional method on earth science students' achievement.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Diego, CA. (ERIC Document Reproduction No. ED418858)

Chapman, J.W., Tunmer, W.E., & Prochnow, J.E. (2000). Early reading-related skills and performance, reading self-concept, and the development of academic self-concept: A longitudinal study*. Journal of Educational Psychology*, 92 (2), 703-708.

Chase, C.I. (1999). *Contemporary assessment for educators*. New York: Longman.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy*. Journal of Child Development*, 70(5), 1098-1120. doi: http://dx.doi.org/10.1111/1467-8624.00081

Chiappetta, E. L., & Koballa, T. R. (2002). *Science instruction in the middle and secondary schools* (5th ed.) Upper Saddle River: Merrill Prentice Hall.

Chonjo, P. N., O-saki, K. M., Possi, M., & Mrutu, M. (1995). *Science education in secondary schools. Situational analysis*. Dar es salaam: Ministry of Education and Culture.

Chonjo, P., & Welford, G. (2001). Reasons for the poor performance of students in A-level science examination in Tanzania. Papers in Education and Development, 21(1), 39-51.

Chow, S, & Yong, S. (2013). Secondary School Students' Motivation and Achievement in Combined Science. *Journal of US-China Education Review B*. 3, (4), 213-228: downloaded from http://files.eric.ed.gov/fulltext/ED542966.pdf

Chumbley, S. B., Haynes, J. C., & Stofer, K. A. (2015). A Measure of Students' Motivation to Learn Science through Agricultural STEM Emphasis. *Journal of Agricultural Education*, 56(4), 107 - 122. doi: 10.5032/jae.2015.04107

Chung, I. (2004). A comparative assessment of constructivist and traditionalist approaches to establishing mathematical connections in learning multiplication. *Education*. 125(2), 271-276.

Cohen L, Manion L, Morrison K (2007). *Research methods in education.* New York, NY: Routledge press.

Coil D, Wenderoth MP, Cunningham M, Dirks C. 2010. Teaching the process of science: Faculty perceptions and an effective methodology. *Journal of CBE Life Sciences Education* 9 (1), 524–535.

Colburn, A. (2000). An inquiry primer. *Journal of Science Scope*, 23(1), 139–140.

Cooley, K. O. (2000). An investigation of the academic self-concept and its relationship to academic achievement in African American college students. *Journal of Black Psychology*, 26 (2), 148-164.

Cortina, J. M. (1993). What is a coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78 (1), 98-104.

Cramer, D. (1997). *Basic statistics for social research*. Routledge. London.

Crawford, B. A. (2007). Learning to teach science in the rough and tumble of practice. *Journal of Research in Science Teaching*, 44(1), 613–642

Crawford, B., (2000). Embracing the Essence of Inquiry: New Roles for Science Teachers. *Journal of Research in Science Teaching*, 37, (9) 916 - 9137.

Crocker, L., and Algina, J. (1986). *Introduction to classical and modern test theory*. Harcourt Brace Jovanovich College Publishers: Philadelphia

Dambudzo, I. (2005). *The relationship between learner self-concept and achievement in secondary schools in Zimbabwe*. (Unpublished Ph.D Thesis), University of South Africa, Zimbabwe. Retrived from *uir.unisa.ac.za/bitstream/handle/10500/2393/01thesis.pdf?sequence=1*

Darus, F, Saat, R. (2014). How do Primary School Students Acquire the Skill of Making Hypothesis. *Malaysian Online Journal of Educational Science*, *Vol 2, Issue 2.* Downloaded from *http://www.moj-es.net/*

Davis L.L. (1992). Instrument review: Getting the most from a panel of experts. *Journal of Applied Nursing Research*, 5, (1)194-197

Deci, E. L., & Ryan, R. M .(2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Journal of American Psychologist,* 55(1), 68-78.

Deci, E.L and Ryan, R. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Journal of Contemporary Educational Psychology,* 25(1), 54–67.

Devellis, R.F. (2011). *Scale development: Theory and applications, (3rd Ed) series 26*. Thousand Oaks, CA: Sage: Retrived from http://www.marketingcenter.de/

Dietz, M. A., & George, K. D. (1970). A test to measure problem solving skills in science of children in grades one, two and three. *Journal of Research in Science Education*, 7 (4), 341 – 351.

Dillashaw, F. G., & Okey, J. R. (1980). Test of integrated process skills for secondary science students. *Journal of Science Education,* 64(5), 601-608.

Dixon, J.K., Adams, T.L. & Hynes, M. (2001). Controlling variables. *Journal of Teaching Children Mathematics,* 8(3), 160-164.

Doolittle, P. E. & Camp, W. G. (1999). Constructivism: The career and technical education perspective. *Journal of Vocational and Technical Education , 16*(1). Retrieved http://scholar.lib.vt.edu/ejournals/JVTE/v16n1/doolittle.html

Dörnyei, Z., & Ottó, I. (1998). Motivation in action: A process model of L2 motivation. *Working Papers in Applied Linguistics*, 4(1), 43–69.

Downing, J., Filer, J., & Chamberlain, R. (1997). Science process skills and attitudes of preservice elementary teachers. Retrieved from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_pdf.

Drayton, B., & Falk, J. K. (2001). Tell-tale signs of the inquiry-oriented classroom. *NASSP Bulletin*, 85(623), 24–34.

DuBay, W. (2004). The principles of readability. Costa Mesa, CA: Impact Information. Retrieved from http://www.impact-information.com/impactinfo/Resources.htm

Duncan, R. G., & Reiser, B. J. (2007). Reasoning across ontologically distinct levels: Students" understandings of molecular genetics. *Journal of Research in Science Teaching,* 44(7), 938-959

Dyer, J. E., Myers, B. E., & Washburn, S. G. (2004). Assessing agriculture teachers' capacity for teaching science integrated process skills*. Journal of Southern Agricultural Education Research*, 54(1), 74-84.

Dyer, J., & Myers, B. (2006). Effects of investigative laboratory instruction on content knowledge and science process skill achievement across learning styles. *Journal of Agricultural Education,* 47(4), 52-63.

Eccles, J.S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Journal of Annual Review of Psychology*, 53, 109–132.

Edelson, D., Gordin, D., & Pea, R. (1999). Addressing the challenge of inquiry-based learning through technology and curriculum design. *Journal of the Learning Sciences*, 8(4), 391-450.

Eisenman, L. (2007). Self-Determination Interventions Building a Foundation for School Completion. *Journal of Remedial and Special Education*, 28 (1), 2–8.

Eiskenkraft, A. (2003). Expanding the 5E Model. *Journal of The Science Teacher*, 70(6), 56-59.

Espinosa, A., A., Monterola, S., L., C, & Punzalan, A., E. (2013). Career-oriented performance tasks in chemistry: Effects on students' integrated science process skills. *Cypriot Journal of Educational Sciences,* 8(2), 211-226.

Fencl, H & Scheel K. (2005). Engaging students: an examination of the effects of teaching strategies on self-efficacy and course climate in a nonmajors physics course*. Journal of College Science Teaching*, 35(1)20–24.

Field, A. (2006). *Discovering statistics using SPSS*, 2nd Ed. London: Sage Publication

Field, S., & Hoffman, A. (1994). Development of a model for self-determination. *Journal of Career Development for Exceptional Individuals*, 17(1), 159-169.

Field, S., & Hoffman, A. (2002). Lessons learned from implementing the Steps to Self- Determination curriculum. *Remedial and Special Education* 23(2), 90-98.

Filson, R. (2001). *In search of real science.* Retrieved from http://www.accessexcellence.com/21st/TL/filson

Fleming, J.S., & Courtney, B.E. (1984). The dimensionality of self-esteem: Hierarchical facet model for revised measurement scales. *Journal of Personality and Social Psychology*, 46(4), 404-421.

Flores, M., Kaylor, M. (2007). The effects of a direct instruction program on the fraction performance of middle school students at-risk for failure in mathematics. *Journal of Instructional Psychology*. 34(2), 84-94.

Forrest, G. M. (1992). Gender differences in school science examinations. *Journal of Studies in Science Education*, 20(4), 87-122.

Foulds, W., & Rowe, J. (1996). The enhancement of science process skills in primary teacher education students. *Australian Journal of Teacher Education,* 21(1). http://dx.doi.org/10.14221/ajte.1996v21n1.2

Freidenreich, H.B., Duncan, R.G., & Shea, N. (2011). Exploring middle school students' understanding of three conceptual models in genetics. *International Journal of Science Education, 33*(17),2323-2349

Frymier, A. (1993). The impact of teacher immediacy on students' motivation: I it the same for all students? *Journal of Communication Quarterly*, 41(4), 454-464

Frymier, A. B., & Shulman, G. M. (1995). What's in it for me? Increasing content relevance to enhance students' motivation. *Journal of Communication Education*, 44(1),40–50.

Frymier, A. B., & Shulman, G.M., & Houser, M. (1996). The development of a learner empowerment measure. *Journal of Communication Education*, 45(5), 181–199.

Furtak, M. (2006). The problem with answers: An exploration of guided scientific inquiry teaching. *Journal of Science Education*, 90 (3), 453-467.

Galabawa, J.C.J (2005). *Enhancing education quality and efficiency: some observation on Kenya, Tanzania and Uganda*. Dar es Salaam: KAD Associates

Gall, M. D., Borg, W. R., & Gall, J. P. (1989). *Educational research: An introduction* (2nd ed.). White Plains, NY: Longman.

Gallagher, S. A. (1994). Middle school classroom predictors of science persistence. *Journal of Research in Science Teaching*, 31(1), 721-734.

Gans, A.M., Kenny, M.e., & Ghany, D.L (2003). Comparing the self-concept of students with and without learning disabilities. *Journal of Learning Disabilities*, 36(3), 285-293.

Gauch, H. G. (2003). *Scientific method in practice*. Cambridge University Press Retrived from HG Gauch - 2003 - books.google.com

Gay, L., & Airasian, P. (2000). *Educational research: Competencies for analysis and application.* Upper Saddle River: Merrill Publishing Company.

George, D., & Mallery, P. (2003). *SPSS for windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Boston: Allyn & Bacon.

Gerardi, S. (2009). Academic self-concept as a predictor of academic success among minority and low-socioeconomic status students. *Journal of College Student Development,* 31(4),402-407.

Germann, P. J., Aram, R., Burke, G. (1996). Identifying patterns and relationships among the responses of seventh-grade students to the science process skill of designing experiments. *Journal of Research in Science Teaching*, 33(1), 79-99.

Ghanem, T. (2003). *The processes of formulatingmhypotheses mand students' difficulties of hypotheses formulation in science learning.* Thesis for the Master Degree of Science Education, Hokkaido University. Retrived from https://www.academia.edu/10156442/

Gibson, S., & Dembo, M. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology,* 76 (4), 569-582.

Gladys, U. (2013). The influence of identified student and school variables on students' science process skills acquisition. *Journal of Education and Practice*, 4(5). Retrived from http://www.iiste.org/Journals/index.php/JEP/article/viewFile/4783/4862

Glasersfeld E. (1995). A constructivist approach to teaching. In: Steffe L. P. & Gale J. (eds.) Constructivism in education. Erlbaum, Hillsdale: 3–15. Available at http://www.vonglasersfeld.com/172

Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science Motivation Questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching*, 46(2), 127-146. Retrieved from https://coe.uga.edu/outreach/programs/science-motivation

Glynn, S. M., Taasoobshirazi, G., & Brickman, P. (2009). Science motivation questionnaire: Construct validation with nonscience majors. *Journal of Research in Science Teaching*, 46(2), 127-146. Retrieved from https://coe.uga.edu/outreach/programs/science-motivation

Glynn, S.M., & Koballa, T.R. Jr., (2006). *Motivation to learn in college science*. In J.J. Mintzes &W.H. Leonard (Eds.), Handbook of college science teaching (pp. 25–32). Arlington,VA: National Science Teachers Association Press.
Retrived from https://coe.uga.edu/assets/files/mse/smqii-10-**glynn-koballa-2006**.pdf

Goldberg, F., and Anderson, J. (1989). Student Difficulties with Graphical Representations of Negative Values of Velocity. *Journal of the Physics Teacher*, 27 (1), 254–60.

Goldberg, M. D., & Cornell, D. G. (1998). The influence of intrinsic motivation and self-concept on academic achievement in second- and third-grade students. *Journal for the Education of the Gifted,* 21(2), 179-205.

Good, T., & Brophy, J. (2000). *Looking in classrooms.* New York: Longman.

Gottfried, A. E. (1990). Academic intrinsic motivation in young elementary school children. *Journal of Educational Psychology,* 82(3), 525-538.

Graham, S., & Harris, K. R. (1989b). Improving learning disabled students' skills at composing essays: Self-instructional strategy training. *Journal of Exceptional Children,* 56(2), 201-214.

Green, Salkind, & Akey (2000). *Using SPSS for windows: Analyzing and understanding data.* London: Prentice Hall.

Greenfield, T. A. (1996). Gender, ethnicity, science achievements and attitudes. *Journal of Research in Science Teaching,* 13(4), 405-412.

Gregg J. (1995). The tensions and contradictions of the school mathematics tradition. *Journal for Research in Mathematics Education*, 26(5), 442-466.

Gregory, R.J. (1992). *Psychological testing: history, principles and applications*, Allyn and Bacon: Boston

Grolnick, W. S., & Ryan, R. M. (1987). Autonomy in children's learning: An experimental and individual difference investigation. *Journal of Personality and Social Psychology,* 52(5), 890–898.

Guthrie, J. T., Wigfield, A., & VonSecker, C. (2000). Effects of integrated instruction on motivation and strategy use in reading. *Journal of Educational Psychology*, 92(2), 331–341.

Guttman, L. (1959). "Introduction to facet design and analysis," in Proceedings of the Fifteenth International Congress of Psychology, Brussels–1957. Amsterdam: North Holland, 130–132.

Hackett, G., & Betz, N. (1989). An exploration of the mathematics self-efficacy/ mathematics performance correspondence. *Journal for Research in Mathematics Education*, 20(2), 261-273.

Hackling, M., & Garnett, P. (1991). Primary and Secondary School Students Attainment of Science Investigation Skills. *Journal of Research in Science Education,* 21 (1), 161-170.

Hadjimarcou, M. I., Constantinou, C. P., & Zacharia, Z. (2009*). Teaching ninth-grade genetics through inquiry*. downloaded from .https://www.researchgate.net/profile/Costas_Constantinou2/publication/281095186_Teaching_NinthGrade_Genetics_Through_Inquiry/links/55d4792c08aec1b042a0d38b.pdf.

Haladyna, T.M. (2004) Developing and validating multiple-choice test items, 3rded. Hillsdale, NJ: Lawrence Erlbaum Associates

Hamachek, D. (1995). Self-concept and school achievement: Interaction dynamics and a tool for assessing the self-concept component. *Journal of Counseling & Development*, 73(4), 419-425.

Hambleton, R & Simon, R. (1980). Steps for Constructing Criterion-Referenced Tests. Laboratory of Psychometric and Evaluative Research Report No. 104. Retrieved from http://files.eric.ed.gov/fulltext/ED189168.pdf

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage Press.

Hambleton, R. K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment*, 10, 229-244.

Hamilton, R., & Swortzel, K. (2007). Assessing Mississippi AEST teachers' capacity for teaching science integrated process skills. Journal of Southern Agricultural Education Research 57(1). Retrived from http://pubs.aged.tamu.edu/jsaer/Vol57Whole.pdf.

Hancock, C., Kaput, J. J., & Goldsmith, L. T. (1992). Authentic inquiry with data: critical barriers to classroom implementation. *Journal of Educational Psychologist*, 27(3), 337-364.

Haque, A., & Sarwat, K. (1998). Age, gender, and achievement effects on academic self-concept of high school children. *Pakistan Journal of Psychological Research*, 13(1), 35-42.

Harlen, W. (1999). Purposes and procedures for assessing science process skills. Assessment in Education. *Journal of Principles, Policy & Practice*, 6(1), 129-144.

Harlen, W. (2000). *The teaching of science in primary schools.* (3rd Ed).London: David Fulton Publishers Ltd.

Harlen, W., and Jelly, S. (1997). *Developing science in the primary classroom* Essex, England: Addison Wesley Longman, Ltd.

Harris, T. L. & Hodges, R. E. (Eds.). (1995). *The literacy dictionary: the vocabulary of reading and writing*. Newark, Delaware: International Reading Association.

Hasse, E. (2003).*Reform in Biology education: Teachers' implementation of a new Biology curriculum*. North Carolina: Unpublished dissertation submitted to the Graduate Faculty, North Carolina State University USA.
Retrived from repository.lib.ncsu.edu/ir/bitstream/1840.16/3318/1/etd.pdf

Hattie, J.A.C. (2012). Visible learning for teachers. Maximizing impact on achievement. Oxford, UK: Routledge.

Haury, D.L. (1993). *Teaching science through inquiry*. ERIC clearinghouse for science mathematics, and environmental education. (ERIC Document Reproduction Service No. ED359048).Retrived from http://www.uhu.es/gaia-inm/invest_escolar/httpdocs/biblioteca_pdf/14_HAURY

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Journal of Psychological Assessment*, 7(2), 238–247.
Henson, R., & Roberts, J. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Journal of Educational and Psychological Measurement,* 66(4), 393-416. doi:10.1177/0013164405282485

Henson, R.K. (2001). Understanding internal consistency reliability estimates: a conceptual primer on coefficient Alpha. *Journal of Measurement and Evaluation in Counseling and Development,* 34(3),177-188.

Hewitt, M. A., & Homan, S. P. (2004). Readability level of standardized test items and student performance: The forgotten validity variable. *Journal of Reading Research and Instruction,* 43(2),1-16.

Hiebert, S. (2007).Teaching simple experimental design to undergraduates: Do your students understand the basics?. *Journal of Advances in Physiology Education* 31(2), 82-92.

Hills, J. R. (1981). *Measurement and evaluation in the classroom* (2nd ed.). Columbus, OH: Charles E. Merrill.

Hinkle DE, Wiersma W, Jurs SG. (2003). *Applied statistics for the behavioral sciences.* 5th ed. Boston: Houghton Mifflin.

Hivey, W., Patterson, H.L., & Page, S.A. (1968). A universe-defined system of arithmetic achievement test. *Journal of Educational Measurement*, 5(1),275-290.

Hofstein, A. & Lunetta, V.N. (2004). The laboratory in science education: Foundations for the 21st century. *Journal of Science Education*, 88(1), 25 -54 .

Hoge, D. R., Smit, E. K., & Crist, J. T. (2012). Reciprocal effects of self-concept and academic achievement in sixth and seventh grade. *Journal of Youth and Adolescence*, 34(1), 295-314.

Hough, L.W., & Piper, M.P. (198%). The relationship between attitude towards science and science achievement. *Journel of Research in Science Teaching*, 19 (I), 33-38.

Houlden, R.L., Raja, J.B., Collier, C.P., Clark, A.F. and Waugh, J.M. (2004) Medical students' perceptions of an undergraduate research elective. *Journal of Medical Teacher*, 26 (7), 659–619.

Huitt, W. (2011). Bloom et al.'s taxonomy of the cognitive domain. *Educational Psychology Interactive*. Valdosta, GA: Valdosta State University. Retrieved from http://www.edpsycinteractive.org/topics/cognition/bloom.html

Huitt, W. (1996). Measurement and evaluation: Criterion- versus norm-referenced testing. *Educational Psychology Interactive*. Valdosta, GA: Valdosta State University. Retrieved from http://www.edpsycinteractive.org/topics/measeval/crnmref.html

Hyman, B and Shephard, A. (1980). The Development of an operational definition, *The Journal of Mind and Behavior*, 1(2), 227-246

Ishumi, A., & Nyirenda, D. (2004). *Philosophy of education. An introduction to concepts, principles and practice*. Dar es Salaam: DUP Ltd

Jansen, M., Schroeders, U., & Lüdtke, O. (2014). Academic self-concept in science: Multidimensionality, relations to achievement measures, and gender differences. *Journal of Learning and individual differences,* 30(1), 11–21. doi: 10.1016/j.lindif.2013.12.003

Jennings, B. (2004). Genetic literacy and citizenship: possibilities for deliberative democratic policymaking in science and medicine. *Journal of the Good Society*, 13 (1), 38–44

John, O.P., & Benet-Martinez, V. (2000). Measurement: Reliability, construct validation, and scale construction. In H. T. Reis & C. M. Judd (Eds.), Handbook of research methods in social psychology, pp. 339-369. New York: Cambridge University Press.

Johnson, J. O. (1996). *Child psychology*. Calabar, Nigeria: Wusen Press Limited.

Johnson, R. L., Penny, J., & Gordon, B. (2000). The relation between score resolution methods and interrater reliability: an empirical study of an analytic scoring rubric. *Journal of Applied Measurement in Education* 13(1), 121-38.

Johnson, S. (1987). Gender differences in science: Parallels in interest, experience, and performance. *International Journal of Science Education*, 9(1), 467-481.

Johnstone A.H. and Al-Shuaili, A. (2001). Learning in the laboratory: some thoughts from the literature. *Journal of University Chemistry Education*, 5(1), 1-10.

Justice, C., Rice, J., Warry, W., Inglis, S., Miller, S., and Sammon, S. (2007). Inquiry in higher education: Reflections and directions on course design and teaching methods. *Journal of Innovative Higher Education,* 31(1), 201-214

Kaplan, R.M. and Saccuzzo, D.P. (2001). *Psychological testing: principle, applications and issues* (5th Edition), Belmont, CA: Wadsworth

Karteroliotis, C., & Gill, D.L. (1987). Temporal changes in psychological and physiological components of state anxiety. *Journal of Sport & Exercise Psychology*, 9(1), 261-274.

Kehoe, J. (1995). Basic item analysis for multiple-choice tests. *Journal of Practical Assessment, Research & Evaluation,* 4(10). Retrieved from http://PAREonline.net/getvn.asp?v=4&n=10.

Keil, C., Haney, J., & Zoffel, J. (2009). Improvements in student achievement and science process skills using environmental health science problem-based learning curricula. *Electronic Journal of Science Education*, 13 (1), 18-31. Retrieved from http://ejse.southwestern.edu.

Keller, J. M. (1983). Motivational design of instruction. In C. M. Reigeluth (Ed.), Instructional design theories: An overview of their current status (pp. 383–434). Hillsdale, NJ: Erlbaum. Retrived from https://sites.google.com/site/motivationataglanceischool/arcs-motivation-model

Kelly, G.A. (1991). *The psychology of personal constructs: A theory of personality.* London: Routledge.

Ketelhut, D. (2007). The impact of student self-efficacy on scientific inquiry skills: An exploratory investigation in river city, a multi-user virtual environment. *Journal of Science Education and Technology*, 16(1). doi: 10.1007/s10956-006-9038-y

Keys, C.W. & Bryan, L.A. (2001). Co-Constructing Inquiry-Based Science with Teachers : Essential Research for Lasting Reform. *Journal of Research in Science Teaching,* 38 (6), 631-645.

Kibga, E. (2004). *The role of practical assessment in teaching and learning of Olevel physics in Tanzania*. M.Ed dissertation. Dar es Salaam, University of Dar es Salaam.

Kincaid, P., Aagard, A., O'Hara, W., Cottrell, K .(1981). Computer Readability Editing System. *IEEE Transactions on Professional Communications*. 24 (1): 38–42.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Journal of Educational Psychologist*, 41(2), 75–86.

Kincaid, JP., Braby, R., Mears, J .(1988). Electronic authoring and delivery of technical information. *Journal of Instructional Development*, 11 (1) 8–13. doi:10.1007/bf02904998.

Klare, G. (1976). A second look at the validity of readality formulas. *Journal of Reading Bebavior*, 8(1), 129- 152

Klein, S. P., Jovanovic, J., Stecher, B. M., McCaffrey, D., Shavelson, R. J., Haertel, E., Solano-Flores, G.,& Comfort, K. (1997). Gender and racial/ethnic differences on performance assessments in science*. Journal of Educational Evaluation and Policy Analysis,* 19(2), 83-97.

Klein, S. P., Kuh, G. D., Chun, M., Hamilton, L., & Shavelson, R. (2005). An approach to measuring cognitive outcomes across higher education institutions. *Journal of Research in Higher Education*, 46(3), 251-276.

Klymkowsky, M., Shi, J., Power, J. (2011). Revealing student thinking about experimental design and the roles of control experiments. *International Journal for the Scholarship of Teaching and Learning*, 5(2). Retrieved from http://www.georgiasouthern.edu/ijsotl

Klymkowsky, M.W. **(**2010)**.** Thinking about the conceptual foundations of the biological sciences. *CBE Life Science Education*, 9(1), 405-417.

Koballa,T., & Crawley, F. (1985). The influence of attitude on science teaching and learning. *Journal of School Science and Mathematics*, 85 (3), 222-232.

Kok-Auntoh, & Woolnough, B. E. (1994). Science process skills: Are they generalisable?. *Journal of Research in Science and Technological Education*, 12(1), 31-42.

Koslowski, B. (1996). Theory and evidence: the development of scientific reasoning. Cambridge, MA: MIT Press.

Krathwohl, D.R. (1993). *Methods of educational and social research: an integrated approach*. New York: Longman

Kremer, B.K., & Walberg, H.J. (1981). A Synthesis of social and psychological influences on sciencelearning. *Journal of Science Education*, 65(1), 11-23

Krzanowski, W. J., & Woods, A. J. (1984). Statistical aspects of reliability in language testing. *Journal of Language Testing,* 1(1), 1-20.

Kyle, W. C., Bonnstetter, R. J., & Gadsden, T., Jr. (1988). An implementation study: An analysis of elementary students' and teachers' attitudes toward science in process-approach vs. traditional science classes. *Journal of Research in Science Teaching*, 25(2), 103–120

Lavigne, G. L., Vallerand, R. J., & Miquelon, P. (2007). A motivational model of persistence in science education: A self determination theory approach. *European Journal of Psychology of Education*, 22(3), 351-369.

Lawrence, D. (1996). *Enhancing self-esteem in the classroom*. London: Paul Chapman.

Lawson, A. E. (2001). Using the learning cycle to teach Biology concepts and reasoning patterns. *Journal of Biological Education*, 35, 65-169.

Lebow D. (1993). Constructivist values for systems design: Five principles toward a new mindset. *Educational Technology Research and Development*, 41(1), 4–16.

Leeuw, H. (2003). *Upgrading of science and mathematics teachers for an educational leadership role in Tanzania.* Master Thesis. Enschede: University of Twente.

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Journal of Review of Educational Research*, 60(1), 1–64.

Lens, W., & Rand, P. (1997). Combining intrinsic goal orientation with professional instrumentality/utility in student motivation. *Polish Psychological Bulletin*, 28(1), 103–123.

Leonard, W. H., Speziale, B. J., and Penick, J. E. (2001).Performance assessment of a standards-based high school biology curriculum. *Journal of American Biology Teacher,* 63(5): 310–316.

Lewis, J., & Wood-Robinson, C. (2000). Genes, chromosomes, cell division and inheritance: Do students see any relationship? *International Journal of Science Education*, 22(1), 177-195.

Lewis, J and Kattman, U. (2004). Traits, genes, particles and information: re-visiting students' understandings of genetics. *International Journal of Science Education,* 26 (2), 195–206

Liang, L. L. & Gabel, D. L. (2005). Effectiveness of a constructivist approach to science instruction for prospective elementary teachers. *International Journal of Science Education, 27*(10), 1143-1162. Retrieved from http://dx.doi.org/10.1080/09500690500069442

Liebert, R.M., & Morris, L.W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Journal of Psychological Reports*, 20(1), 975-978.

Linn, M. C., Clement, C., & Pulos, S. (1983). Is it formal if it's not physics? (The influence of content on formal reasoning). Journal of Research in Science Teaching, 20(8), 755–770.

Lin, Y., McKeachie, W. J., & Kim, Y. C., College Student Intrinsic and/or Extrinsic Motivation and Learning. *Learning and Individual Differences*, 2001. 13(3): p. 251-258.

Lock, R. (1993). Assessment of practical skills. Part 2. Context dependency and construct validity. *Research in Science and Technological Education*, 8 (1), 35-52.

Longo, C. M. (2011). Designing inquiry oriented science lab activities. *Middle School Journal*, 43 (1) 6-15.

Lord, T., & Orkwiszewski, T. (2006). Moving from didactic to inquiry-based instruction in a science laboratory. *Journal of the American Biology Teacher*, 68(1), 342-345.

Lord, F.M. & Novick, M.R. (1968). Statistical theories of mental test scores.Reading, MA: Addison-Wesley.

Lunetta, V. N., Hofstein, A. and Clough, M. P. (2007*). Teaching and learning in the school science laboratory. An analysis of research, theory, and practice*. In, S. K. Abell and N. G. Lederman (Eds), Handbook of Research on Science Education (pp. 393–431). Mahwah, NJ: Lawrence Erlbaum Associates.

Luo, M., & Dappen, L.(2005). Mixed-methods design for an objective-based evaluation of a magnet school assistance project. *Evaluation and Program*

Planning, (28)1, 109-118, ISSN 0149-7189, Retrieved from (http://0-www.sciencedirect.com.uncclc.coast.uncwil.edu/science/article/pii/S0149718904000849)

Lynn, M. R. (1986). Determination and quantification of content validity. *Journal of Nursing Research*, 35(1),382—385.

Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education*, 30, 520–541.

Mafumiko, F. M. S. (1998). *The role of practical work in chemistry education in Tanzania: Exploration of current practices and potential alternatives*. Master Thesis. Enschede: University of Twente.

Marsh, H. W., Byrne, B. M., & Shavelson, R. (1988). A multifaceted academic self-concept: Its hierarchical structure *and* its relation to academic achievement. *Journal of Educational Psychology,* 80(1), 366-380.

Marsh, H.W., Trautwein, U., Ldtke, O., Kller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Journal of Child Development,* 76(1), 397–416.

Marshall, J. C. & Horton, R. M. (2011). The relationship of teacher-facilitated, inquiry-based instruction to student higher-order thinking. *Journal of School Science and Mathematics, 111*(3), 93-101.
Martin, D. W. (1996). *Doing psychology experiments*. (4th ed.). Pacific Grove, CA: Brooks/Cole.

Marx, J.G., Honeycutt, K.A., Clayton, S.R., & Moreno, N.P. (2006). The Elizabeth Towns Incident: An ınquiry-based approach to learning anatomy developed through high school-university collaboration. *Journal of the American Biology Teacher,* 68(3), 140-147

Mattheis, F. E., & Nakayama, G. (1988). Development of the performance of process skills (POPS) test for middle grades skills. (ERIC Document Reproduction Service No. ED 305 252). Retrived from https://archive.org/stream/ERIC_ED305252/ERIC_ED305252_djvu.txt

Mayer, R. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *Journal of American Psychologist,* 59(1), 14–19.

McKenzie, D and Padilla, M. (1986). The construction and validation of the “Test of Graphing in Science (TOGS)”. *Journal of Research in Science Teaching*, 23(1), 571–579

McMillan, J., & Schumacher, S. (1997). *Research in education: A conceptual introduction*. New York: Harper Collins.

Mechling, K., Bires, N., Kepler, L., Oliver, D., and Smith, B. (1985). *A recommended science competency continuum*.
Retrieved from http://www.maisk6scienceinquiry.org/

Meeker, M, Meeker, R, & Roid, (1985). Structure of intellect leaning ability tests. Los Angeles, CA: Western Psychological Services.

Messick, S. (1980). Test validity and the ethics of assessment. *Journal of American Psychologist*, 35(1), 1012-1027. doi: 10.1037//0003-066X.35.11.1012

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performance as scientific inquiry into score meaning. *Journal of American Psychologist*, 50, 741-749.

Millar, R., and Driver, R. (1987). Beyond Processes. *Journal of Studies in Science Education*, 14, 33-62.

Miller, J. (1998). The measurement of civic scientific literacy. *Journal of Public Understanding of Science,* 7 (1), pp. 203–223

Miller, S. K. (2003). A comparison of student outcomes following problem based learning instruction versus traditional lecture learning in a graduate pharmacology course. *Journal of the American Academy of Nurse Practitioners* , 15(12), 550-556.

Ministrell, J., & van Zee, E. H. (2000). Inquiring into Inquiry Learning and Teaching in Science. Washington, DC: American Association for the Advancement of Science. Retrived from http://www.aaas.org/sites/default/files/migrate/uploads/InquiryFM.pdf

Molitor, L.L., & George, K.D. (1976). Development of a test of science process skills. *Journal of Research in Science Teaching*, 13(5), 405 – 412.

Moreno, R. (2004). Decreasing cognitive load in novice students: Effects of explanatory versus corrective feedback in discovery-based multimedia. *Journal of Instructional Science*, 32, 99–113.

Morogoro Municipal Council. (2014). *Morogoro municipal council:* The profile 2014 Morogoro: Unpublished document

Morris, W.; Davis, A.; Hutchings, H. (1981). Cognitive and emotional components of anxiety: Literature review and a revised worry–emotionality scale. *Journal of Educational Psychology*, 73(4), 541-555. http://dx.doi.org/10.1037/0022-0663.73.4.541

Moskal, B.M., & Leydens, J.A. (2000). Scoring rubric development: Validity and reliability. *Journal of Practical Assessment, Research & Evaluation,* 7(10). Retrived from: http://pareonline.net/getvn.asp?v=7&n=10]

Moss, P.(1998). The role of consequences in validity theory. *Journal of Educational Measurement*, 17(2), 6-12

Mungandi, K. M. (2005).*Development and validation of a test of integrated science process skills for the further education and training learners*. Unpublished dissertation report: University of Pretoria, South Africa

Mushi, P.D. (1992). *Technology: The perception of teachers and students and its place in the secondary school science curriculum.* Unpublished M.Ed.Science Dissertation. University of Leeds, United Kingdom.

Napier, J.D., & Riley, J.P. (1985). Relationship between affective determinants and achievement in science for seventeen-year-olds. *Journal of Research in Science Teaching,* 22(4), 365–383.

Nath, B. (2015).Constructivist approach way to promote self-concept and achievement in science of upper primary students. *International Journal of Multidisciplinary Research and Development,* 2(11), 616-621
Retrived from www.allsubjectjournal.com

Ndalichako, J. L., & Rogers, W. T. (1997). Comparison of finite state score theory, classical test theory, and item response theory in scoring multiple-choice items. *Journal of Educational and Psychological Measurement*, 57(1), 580-589.

Nevo, B. (1985). Face validity revisited. *Journal of Educational Measurement*, 22(4), 287-293.

Nitko, A. J. (1996). *Educational assessment of students, 2nd ed*. Englewood Cliffs, NJ: Prentice-Hall.

Njoroge, G.N., Changeiywo, J.M & Ndirangu, M. (20140. Effects of inquiry-based teaching approach on Secondary School Students' achievement and motivation in Physics in Nyeri county, Kenya. *Academic Research Journals*, 2(1), 1-16

Noss, R., Pozzi, S., and C. Hoyles. (1999). Touching epistemologies: Meanings of average and variation in nursing practice. *Educational Studies in Mathematics*, 40(1), 25–51.

Nugent,G and Kunz, G. (2008). The impact of a field-based, inquiry-focused model of instruction on preservice teachers' science learning and attitudes. *The Electronic Journal of Science Education,* 12 (2), 1-18.
Retrieved from http://ejse.southwestern.edu

Nunnally, J.C., and Bernstein, I.H. (1994) *Psychometric theory,* (3rd edn). Mcgraw-Hill: New York

Nwosu, A.A. & Okeke, E.A.C. (1995). The effect of teacher sensitization of students' acquisition of science process skills. *Journal of Science Teacher Association Nigeria*, 30(1), 39-45.

Ojo, E., & Sola, A. (2007). Effects of project, inquiry and lecture-demonstration teaching methods on senior secondary students' achievement in separation of mixtures practical test. *Journal of Educational and Review*, 2 (6), 124-132.

Olatoye, R. A., & Afuwape, M. O. (2003). Test anxiety as a determinant of examination misdemeanor among some Nigerian secondary school students. *Ibadan Journal of Educational Studies*, 3(182), 32-39.

Oloruntegbe, K.O. & Omoifo C.N. (2000) Assessing process skills in stm education: going beyond paper and pencil tests. *Educational Thought*, 1(1) 35-44

Olson, S., & Loucks-Horsley, S. (Eds.) (2000). Inquiry and the national science education standards: A guide for teaching and learning. Washington, DC: National Research Council. Retrived from https://www.nap.edu/login.php?record_id=9596&page=http%3A%2F%2Fwww.nap.edu%2Fdownload%2F9596

Onwu, G. & Stoffels, N. (2005). Instructional functions in large, under-resourced science classes: Perspectives of South African teachers. *Perspectives in Education,* 23(3), 79-91.

Onwu, G.O., & Mozube, B. (1992). Development and validation of science process skills test for secondary science students. *Journal of Science Teacher Association of Nigeria*, 27(2), 37-43.

Osaki, K. M., & Njabili, A. (2004). *Secondary education sector analysis*. Research Report submitted to Ministry of Science, Technology and Higher education & World Bank. Dar es Salaam: World Bank.

Osaki, K.M. (2000). *The science education in secondary school*. Internal project evaluation. Dar es Salaam: Ministry of Education and Culture.

Osaki, K.M. (2007). Science and mathematics teacher preparation in Tanzania: Lessons from teacher improvement projects in Tanzania since 1965-2006. *Journal of International Educational Cooperation*, 2 (1), 51-64.

Osborne, J., & Collins, S. (2001). Pupils' views of the role and value of the science curriculum: A focus group study. *International Journal of Science Education*, 23(5), 441-467.

Osborne, J; Simon, S & Collins, S. (2003). Attitudes towards science: a review of the literature and its implications. *International Journal of Science Education*, 25(9), 1049-1079.

Ostlund, K. (1998). What research says about science process skills. *Journal of Science Education,* 2 (4).
Retrieved from http://unr.edu/homepage/jcannon/ejse/ejsev2n4

Padilla, J. M., McKenzie, L. D., & Shaw, L. E. (1986). An Examination of line graphing ability of students in grades seven through twelve. *Journal of Scholl Science and Mathematics*, 86 (1), 20-26.

Padilla, M. J. (1990), Science Process Skills- Research Matters - to the Science Teacher. (ERIC Document Reproduction Service No. ED266961). Retrieved from: http://www.narst.org/publications/research/skill.cfm.

Padilla, M. J., Okey, J.R., & Dillashaw, F. G. (1983). The relationship between science process skill and formal thinking abilities. *Journal of Research in Science Teaching,* 20 (3), 239-246.

Painter, J. (2011). *Autonomy, competence, and intrinsic motivation in science education: A self- determination theory perspective*, Ph.D. Dissertation, The University of North Carolina at Chapel Hill: Document URL http://search.proquest.com/docview/1322253395?accountid=15158

Pajares, F. (2001). Self-efficacy beliefs in academic settings. *Journal of Review of Educational Research*, 66(1), 543-578.

Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. *Journal of Reading& writing Quarterly*, 19 (2), 139-158.

Pajares, F., & Valiante, G. (1997). Influence of writing self-efficacy beliefs on the writing performance of upper elementary students. *Journal of Educational Research,* 90, 353-360.

Pajares, F., & Schunk, D. H. (2001). *Self-beliefs and school success: Self-efficacy, self-concept, and school achievement.* In R. Riding & S. Rayner (Eds.), *Self-perception* (pp. 239-266). London: Ablex Publishing.

Peek, L. E., Winking, C., &Peek, G. S. (1995). Cooperative learning activities: Managerial accounting. *Issues in Accounting Education*, 10(1), 111-125.

Pinar, D., & Ceren, T. (2008). Promoting students' learning in genetics with the learning cycle. *The Journal of Experimental Education*, 76, (3), 259-280 URL: http://www.jstor.org/stable/20157486

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33-40.

Pintrich, P.R., & Schunk, D.H. (1996). *Motivation in education: Theory, research and applications* (2nd ed.). Englewood Cliffs, NJ: Merrill Company.

Plowden (1967) Children and their Primary Schools. Report of the Central Advisory Council for Education (England). London: *Retrived from* www.educationengland.org.uk/documents/**plowden/plowden1967**-1.html

Plowright, D. and M. Watkins (2004). There are no problems to be solved, only inquiries to be made, in social work education. *Innovations in Education and Teaching International,* 41, 185-206.

Poon, C., Tan, D., & Tan, A. (2009). Classroom management and inquiry-based learning: Finding the balance. *Journal of Science Scope*, 32 (9), 18-21.

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

Popham, J. W. (1975). *Educational evaluation*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Popham, W.J., & Husek, T.R. (1969). Implications of criterion-referenced measurement. *Journal of Educational Measurement*, (1)6, 1-9.

Pratt, H. & Hackett, J. (1998). Teaching science: The inquiry approach. *Journal of Principal*, 78(2), 2-20.

Prawat, R. S., & Floden, R. E. (1994). Philosophical perspectives on constructivist views of learning. *Journal Educational Psychology*, 29(1), 37–48.

Preece, F. W., & Brotherton, P. N. (1997). Teaching science process skills: Long-term effects on science achievement. *International Journal of Science Education, 19*(8), 895-901.

Prince, M., and R.M. Felder. (2006). Inductive teaching and learning methods: Definitions, comparisons, and research bases*. Journal of Engineering Education* 95 (2): 123–38.

Putwain, D. W. (2007). Test anxiety in UK schoolchildren: Prevalence and demographic patterns. *British Journal of Educational Psychology*, 77, 579–593.

Radford, D. (1988). Integrating process skills instruction into the traditional science curriculum. Paper presented at the annual meeting of the National Association forResearch in Science Teaching, Lake Ozark, MO. (ERIC document ED 291 588). Retrived from http://files.eric.ed.gov/fulltext/ED354145.pdf

Rambuda, A., & Fraser, W. (2004). Perceptions of teachers of the application of science process skills in the teaching of Geography in secondary schools in the Free State province. *South African Journal of Education*, 24(1), 10 – 17.

Rambuda,A & WJ Fraser, W. (2006). Perceptions of teachers of the application of science process skills in the teaching of Geography in secondary schools in the Free State province. *South African Journal of Education*, 24 (1), 10-17

Ramsey, J. (1993). Developing conceptual storylines with the learning cycle. *Journal of Elementary Science Education*, 5(2), 1-20.

Randolph, J. (2008). Multidisciplinary Methods in Educational Technology Research and Development. Julkaisuja. Hameenlina, Finland.
Retrived from http://justusrandolph.net/articles/multidisciplinary_methods.pdf

Reckase, M. D. (1998). Consequential validity from the test developer's perspective. *Journal of Educational Measurement*, 17 (2), 13-16.

Reese, E., Bird, A., & Tripp, G. (2007). Children's self-esteem and moral self: Links to parent child conversations regarding emotion. *Journal of Social Development*, 16(1), 460-478.

Reeve, J., Nix, G., & Hamm, D. (2003). Testing models of the experience of self-determination in intrinsic motivation and the conundrum of choice. *Journal of Educational Psychology,* 95(1), 375-392.

Rehorek J.S. (2004). Inquiry-Based Teaching: An example of descriptive science in action. *Journal of American Biology Teacher*, 66(7), 493-500.

Reinhart, C.S. & Rallis, S.F. (1994). *The qualitative-quantitative debate: New perspectives.* San Francisco: Jossey-Bass.

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences*, 13(1), 273–304.
Renner, J. W., Abraham, M. R., & Birnie, H. H. (1988). The necessities of each phase of the learning cycle in teaching high school physics. *Journal of Research in Science Teaching,* 25, 39-58.

Rezba, R.J., Sparague, C.S., Fiel, R.L., Funk, H.J., Okey, J.R., & Jaus, H.H. (1995). *Learning and assessing science processes*. Dubuque: Kendall/Hunt Publishing Company

Richards, J.C., Platt, J., and Platt, H. (1992). *Dictionary of language teaching and applied linguistics*. London: Longman.

Riggs, I. (1991). Gender Differences in Elementary Science Teacher Self-Efficacy. ERIC residue No ED340705. Retrieved from http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=ED340705.

Riggs, I., & Enochs, L. (1990). Towards the development of an elementary teacher's science teaching efficacy belief instrument. *Journal of Science Education*, 74(2), 625-637.

Rissing S.W., & Cogan J.G. (2009). Can an ınquiry approach ımprove college student learning in a teaching laboratory? *Journal of CBE Life Science Education*, 8(1), 55–61.

Roid, G.H. & Haladyna, T.M. (1982). A technology for test-item writing. New York: Academy Press.

Rolf, G., Peter K, & Wiltrud, W, eds (2010). Taking Part in Democracy: EDC/HRE Lesson Plans for Upper Secondary Level. Retreaved from http://www.living-democracy.com/textbooks/volume-4/introduction-2/

Ross, A .(1988) .Controlling variables: A meta-analysis of training studies. *Journal of Review of Educational Research Winter* , Vol. 58, No. 4, pp. 405-437

Ross, J.A. (1990). Learning to control variables. Main effects and aptitude treatment interactions of two rule-governed approaches to instruction. *Journal of Research in Science Teaching*, 27(6) 523-539.

Ross, K. (2005) (Ed). Educational research: some basic concepts and terminology. Paris: UNESCO International Institute for Educational Planning. Retrieved from http://www.unesco.org/iiep on 10th June 2013.

Roth. W. M., & Roychounhury, A. (1993). Development of science process skills in authentic contexts. *Journal of Research in Science Teaching,* 30 (2), 127-152.

Rowe, J., & Foulds, W. (1996).The enhancement of science process skills in teacher education students. *Australian Teacher Education*, 21(1), 16-22.

Rowland, P. M. (1990). Using science activities to internalize locus of control and influence attitudes towards science. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching (63rd, Atlanta, GA, April 8-11, 1990). (ERIC Document Reproduction Service No. ED 325 333)
Retrived from http://eric.ed.gov/?id=ED325333

Rubin, R. L., & Norman, J. T. (1992). Systematic modeling versus the learning cycle: Comparative effects on integrated science process skill achievement. *Journal of Research in Science Teaching, 29*(7), 715-727.

Rudner, L.M., & Schafer, W.D. (2001). Reliability. ERIC Database. Retrieved from http://www.eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED458213.

Ruiz-Primo, M. A. & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, 33(6), 569-600.

Rutherford, F. J. (1993). Hands-on: A means to an end. 2061 Today, 3(1), 5.
Retrieved from
http://www.project2061.org/publications/2061Connections/archive.htm.

Sanchez, Z., & Betkouski, M. (1986).A study of factors affecting student performance in community college general chemistry courses. San Francisco: National Association for Research in Science. Retrieved from http://eric.ed.gov/?id=ED275549

Sarason, I. G., Sarason, B. R., & Pierce, G. R. (1990). Anxiety, cognitive interference, and performance. *Social Behavior and Personality*, 5, 1–18.

Scharmann, L. C. (1989). Development of science process skill instruction. *Journal of Research in Science Teaching*, 26 (8), 715-726.

Schindler, V. (2014). Chapter 10. An Analysis of the Effectiveness of the Intervention. In Occupational Therapy in Forensic Psychiatry (pp152–170). Abingdon, Oxon: Routledge. Retrived from
http://www.tandfonline.com/doi/abs/10.1300/J004v20n03_10?journalCode=womh20
Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(1), 262-274.

Shock, S.A. & Coscarelli, W.C. (2000). Criterion-Referenced Test Development (2nd ed.). Silver Springs, MD: International Society for Performance Improvement.

Schulz, L. E., & Gopnik, A. (2004). Causal learning across domains. *Journal of Developmental Psychology*, 40(2), 162–176.

Schunk, D. H., & Swartz, C. W. (1993). Goals and progress feedback: Effects on self-efficacy and writing achievement. *Journal of Contemporary Educational Psychology,* 18(1), 337-354

Schunk, D.H., Pintrich, P.R., & Meece, J.L. (2008). Motivation in education (3rd ed.). Upper Saddle River, NJ: Pearson.

Schunk, D. H. (2005). Self-regulated learning: The educational legacy of Paul R. Pintrich.*Educational Psychologist, 40*, 85-94.

Schunk, D. H., & Usher, E. L. (2012). Social cognitive theory and motivation. In R. Ryan (Ed.). The Oxford handbook of human motivation (pp. 13-27). New York, NY: Oxford University Press. DOI: 10.1093/oxfordhb/9780195399820.013.0002

Schweinhart, L. J., Weikart, D.P., Larner, M. B. (1986) Consequences of three pre-school curriculum models. *Early Childhood Research Quarterly*, 1(1), 15-45.

Schwichow, M., Croker, S, Zimmerman, C., Höffler, T, & Härtig, H. (2016). Teaching the control-of-variables strategy: A meta-analysis. *Journal of Developmental Review*, 39 (2) 37–63

Sefu, M & Siwale, E. (1977). *The development of primary education in Tanzania*. Ontario: Brock University

Seldin, P. (1999). Changing Practices in Evaluating Teaching. Bolton, Mass: Anker. Retrieved from http://www.iucat.iu.edu/iuk/3688627

Seligman, M. E. P. & Csikszentmihalyi, M. (2000). Positive Psychology: An introduction. *Journal of American Psychologist*, 55(1), 5-14.

Semali, M &Mehta, K. (2012). Science education in Tanzania: Challenges and policy responses. *International Journal of Educational Research*, 53(1), 225–239

Seok, Oh. (2010). How can teachers help students formulate scientific hypotheses? Some strategies found in abductive inquiry activities of earth science. *International Journal of Science Education*, 32(4), 541-560 doi:10.1080/09500690903104457

Shadmi, Y. (1981). Teaching 'control of variables' to primary school teachers. *Journal of Physics Education*, 16, 93-98

Shahali, E & Halim, L. (2010). Development and validation of a test of integrated science process skills. *Procedia Social and Behavioral Sciences*, 9(1), 142-146.

Shaibu, A.M., & Mari, J.S. (1997). Gender-related difference in the understanding of science process skills amongst secondary school students in some Nigerian schools. *Journal of Science Teachers Association of Nigeria*, 32(2), 21-27.

Shavelson, R.J., Baxter, G.P., & Pine, J. (1992). Performance assessments: Political rhetoric and measurement reality. *Journal of Educational Researcher*, 21(4), 22-27.

Shavelson, R.J; Baxter, G. P & Pine, J. (1999). Performance assessment in science. *Journal of Applied Measurement in Education*, 4 (4) 347-362

Shayer, M., & Adey, P. (1981). Towards a science of science teaching. London: Heinemann
Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/tea.3660200411/pdf

Shemwelekwa, R. (2008). *The effectiveness of adoption of competence based education for teaching and learning mathematics in secondary schools. Dar es Salaam.* Unpublished Master Dissertation, University of Dar es Salaam, Tanzania.

Sherry, L. (1997). Item analysis. Retrieved from http://carbon.cudenver.edu/~lsherry/rem/ item analysis.html

Shultz, K., & Whitney, D. (2005). *Measurement theory in action: Case studies and exercises.* Thousand Oaks, CA: Sage Publications.

Shuttleworth, M (2009). Between subjects design. Retrieved May 10, 2016 from Explorable.com: https://explorable.com/between-subjects-design

Shymansky, J. A., & Penick, J. E. (1981). Teacher behavior does make a difference in hands- on science classrooms. *School Science and Mathematics*, 81(5), 412-22.

Shymansky, J., L. Hedges, and G. Woodworth. (1990). A reassessment of the effects of inquiry-based science curricula of the 60's on student performance. *Journal of Research in Science Teaching*, 27 (2), 127–44.

Simon, M.S., & Zimmerman, J.M. (1990). Science and writing. *Journal of Science and Children*, 18 (3), 7-18.

Singamaneni,S (2011) *Exploring content validity, item level analysis and predictive validity for two algebra progress monitoring measures*: Unpublished Thesis submitted to the graduate faculty in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE, Iowa State University

Sireci, S.G. (1998). Gathering and analyzing content validity data. *Journal of Educational Measurement,* 5(1), 299-321.

Skinner, BF (1974). About Behaviorism . Random House , NY, USA . Retrieved from http://fitelson.org/prosem/skinner_2.pdf

Sodian B., Zoitchek D. & Carey, S. (1991). Young children's differentiation of hypothetical beliefs from evidence. *Journal of Child Development*, 62(1), 753-766.

Souchek, R., & Meier, M. (1997). Teaching information literacy and scientific process skills: An integrated approach*. Journal of College Teaching*, 45(4), 128-31

Specht, S. M. (2004). *Bucket o' constructs: Introduction to operational definitions and technical writing.* Poster session presented at the Finding out best practices in teaching research methods in psychology conference, Atlanta, GA. Retrived from http://www.jolley-mitchell.com/teachRDE/C05/C5OperationalDefn.html

Spector,P.E. (1992). *Summated rating scale construction*. Newbury Park, CA: SAGE Publications, Inc. doi: http://dx.doi.org/10.4135/9781412986038

Spronken-Smith, R., Angelo, T. Matthews, H, O'Steen, B. & Robertson, J. (2007). How effective is inquiry-based learning in linking teaching and research? An International Colloquium on Policies and Practices, Winchester. Retrieved from: http://portal live.solent.ac.uk/university/rtconference/colloquium_papers.aspx

Stephens, S. (2000). *All about readability*. Retrived July 13 2014 from http://www.plainlanguagenetwork.org/stephens/readability.html

Stiggins, R. J. (1987). The design and development of performance assessments. *Educational Measurement, Issues and Practice*, 6(1), 33-42.

Stipek, D. (1998). *Motivation to learn: From theory to practice.* Boston: Allyn and Bacon

Streiner, D. L. (2003).Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, 80(1), 99–103.

Sweller, J. (1994). Cognitive load theory, learning difficulty and instructional design, *Journal of Learning and Instruction* 4(1)295–312.

Tan, M & Temiz, B.K. (2003). The importance and role of the science process skills in science teching. *Pamukkale University Journal of Faculty of Education*, 13(1), 89-101

Taylor, N and Corrigan, G. (2005).Empowerment and confidence: Pre-service teachers learning to teach science through a program of self regulated learning. *Canadian Journal of Science, 5 (1), 41–61*

Thompson, J. and Soyibo, K. (2001). Effects of lecture, teacher demonstrations, discussion and practical work on 10th grader's attitude to chemistry and understanding of electrolysis. *Research in Science and Technological Education*, 20, 25–37.

Tiemann, P.W., & Markle, S.M. (1983). *Analyzing instructional content: A guide to instruction and evaluation (2nd ed.).* Champaign, IL: Stipes.

Tilya, F., & Mafumiko, F. (2008).*Compatibility of teaching methods and competence-based curriculum in Tanzania*. Dar es Salaam: Unpublished Paper

Tobin, K. G. and W. Capie (1982). Development and validation of a group test of integrated science processes. *Research in Science Teaching*, 19(1), 133-141.

Traub, R. (1997). Classical Test Theory in Historical Perspective*. Journal of Educational Measurement: Issues and Practice,* 16 (4), 8-14. Doi: 10.1111/j.1745-3992.1997.tb00603.x

Trautwein, U., Liidtke, 0., Koller, 0., & Baumert, J. (2006). Self-esteem, academic selfconcept, and achievement: How environment moderates the dynamics of self-concept. *Journal of Personality and Social Psychology*, 90, 334-349.

Trigwell, K. (1990). The effects of an alternative science degree programme on the participation of women in the physical sciences. *International Journal of Science Education*, 12(1), 25-40.

Tsui, C and Treagust, D. (2010). Evaluating secondary students' scientific reasoning in genetics using a two-tier diagnostic instrument. International Journal of Science Education, 32 (8) (2010), pp. 1073–1098

Tuan, H. L., Chin, C. C., & Shieh, S. H. (2005). The development of a questionnaire to measure students' motivation towards science learning. International Journal of Science Education, 27(6), 639–654.

Turaib, H. (2015). Assessing Students' Understanding of Control of Variables across Three Grade Levels and Gender. Journal of International Education Studies,9(1): Retrieved from URL: http://dx.doi.org/10.5539/ies.v9n1p44

Turner, J & Oakes, P (1986). The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence. *British Journal of Social Psychology*. 25 (3), 237–252.

Tyler, R. W. (1949). Basic principles of curriculum and instruction. Chicago: University of Chicago Press. ISBN 0-226-82031-9. Retrieved from http://blogs.ubc.ca/ewayne/files/2009/02/tyler_001.pdf

United Republic of Tanzania. (2005). *Ordinary secondary level biology syllabus*. Dar es Salaam: Tanzania Institute of Education.

United Republic of Tanzania. (2010). *Advanced secondary level biology syllabus*. Dar es Salaam: Tanzania Institute of Education.

United Republic of Tanzania. (2008). *The development of education national report of the United Republic of Tanzania*. Dar es Salaam: Ministry of Education and Vocational Training.

Vallerand, R.J., & Bissonnette, R. (1992). Intrinsic, extrinsic, and amotivational styles as predictors of behavior: A prospective study. *Journal of Personality*, 60(1), 599–620.

VeisiKahre S, Imani S, Yosef Zade MR, VeisiPour M, Moradhaseli Y, Amiri RK. Effectiveness of problem solving Training on self concept academic high school students in holillan kahreh. *International Journal of Educational and Psychological Researches*, 1(2),131-134
Available from: http://www.ijeprjournal.org/text.asp?2015/1/2/131/152227

Voorhees, C.M., Brady, M.K., Calantone, R., Ramirez, E. (2015). Discriminant validity testing in marketing: an analysis, causes for concern, and proposed remedies. *Journal of the Academy of Marketing Science*, 1(3),1–16

Wang, M. C., Haertel, G., & Walberg, H. J. (1993). Toward a knowledge base of school learning. *Journal of Review of Educational Research*, 73(3), 249-294.

Ware, S .(1992). *The education of secondary science teachers in developing countries*. Washington, DC: World Bank. Retrieved from http://siteresources.worldbank.org/EDUCATION/Resources/278200-1099079877269/547664109908063795/The_education_of_secondary_science_teachers_EN92.pdf

Warner J. A & Myers, B.E. (2008). Implementing inquiry based teaching methods. Retrieved from: http://edis.ifas.ufl.edu/pdffiles/WC/WC07600.pdf.

Watson, R., Prieto, T. & Dillon, J.S. (1995). The effect of practical work on students' understanding of combustion. *Journal of Research in Science Teaching*, 32(5), 487–502.

Webster, C ; Mindrila, D and Glenn Weaver,D. (2011). The influence of state motivation, content relevance and affective learning on high school students' intentions to use content following completion of compulsory physical education. *Journal of Teaching in Physical Education*, 30(1), 231-247

Weinburgh,M.H and Englehard,G. (1994). Gender, prior academic performance and beliefs as predictors of attitudes towards biology laboratory experiences, School Science & Mathematics, 94 (3) (1994), pp. 118–123

Westbrook, S.L and Rogers, L.N. (1994). Examining the development of scientific reasoning in ninth-grade physical science students. *Journal of Research in Science Teaching,* 31(1), 65–76

Wetzel, D. (2008). Problem solving and science process skills: Science investigation skills. Retrieved from http://www.suite101.com

Wiggins, G. (1989). A true test: toward more authentic and equitable assessment. *Journal of Phi Delta Kappan*, 70 (9), 703-713.

Wiggins, G. P. (1998). *Educative assessment: Designing assessments to inform and improve student performance.* San Francisco: Jossey-Bass Publishers.

Wolters, C., & Rosenthal, H. (2000). The relation between students' motivational beliefs and attitudes and their use of motivational regulation strategies. *International Journal of Educational Research*, 33, 801 – 820.

Worthington, R., & Whittaker, T. (2006). Scale development research: A content analysis and recommendations for best practices. *Journal of Counseling Psychologist,* 34, 806-838. doi:10.1177/0011000006288127

Wright, J. D. (1982). The effect of reduced readability text materials on comprehension and biology achievement. *Journal of Science Education*, 66, 3-13.

Wright, S., Sandra, P., & William, L. (1997). Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation. *Journal of Personnel Evaluation in Education*, 11(2), 57-67

Yager, R. (1991). The constructivist learning model. *Journal of the Science Teacher*, 58(6), 53-57.

Yager, R., & Lutz, M. (1994). Integrated science: The importance of "how" verses "what." *Journal of School Science and Mathematics*, 94 (1), 338-346.

Yen, W. M. (1998). Investigating the consequential aspects of validity: Who is responsible and what should they do?.*Educational Measurement*, 17 (2), 5-6.

Zeidner, M., & Schleyer, E. (1999). The big-fish-little-pond effect for academic self-concept, test anxiety, and school grades in gifted children. *Journal of Contemporary Educational Psychology,* 24, 305–329.

Zeitler,W.R. (1981). The influence of the type of practice in acquiring process skills. *Journal of Research in Science Teaching,* 18 (3), 188-191

Ziman, J. (2000). Real Science. Cambridge: Cambridge University Press. Retrieved from http://catdir.loc.gov/catdir/samples/cam032/99038905.pdf

Zimmerman, B., & Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment. *American Education Research Journal*, 31(6), 8457862.

Zimmerman, C. (2000). The development of scientific reasoning skills. *Journal of Developmental Review*, 20(2), 99-149.

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Journal of Developmental Review*, 27(1), 172-223

Zoller, U. (1991). Teaching/learning styles, performance, and students' teaching evaluation in S/T/E/S-focused science teacher education: A quasi-quantitative probe of a case study.*Journal of Research in Science Teaching*, 28(7), 593–607.

Appendix I: Biology Process Skills Test (BPST)

BIOLOGY PROCESS SKILLS TEST (BPST)

INSTRUCTIONS

DURATION: 60 minutes

i. DO NOT WRITE ANYTHING IN THIS PAPER
ii. ANSWER ALL THE QUESTIONS ON THE ANSWER GRID PROVIDED, BY PUTTING A CORRECT ALTERNATIVE OF YOUR CHOICE.
iii. PLEASE DO NOT GIVE MORE THAN ONE ANSWER PER QUESTION.

1. In the Biology Lab Zai has an aquarium in which she keeps goldfish for her student's practical work. She notices that the fish are very active sometimes but not at others. She wonders what affects the activity of the fish and decided to conduct an experiment to find out the reason. What is a hypothesis she could test about factors that affect the activity of the goldfish in the Aquarium?

A. The more you feed fish, the larger the fish become.
B. The more active the fish, the more food they need.
C. The more oxygen in the water, the larger the fish become.
D. The more light on the aquarium, the more active the fish.

2. A learner observed that anthills (termite moulds) in a certain nature reserve tend to lean towards the west, instead of being straight. In this area, the wind blows towards the direction in which the anthills lean. Which of the following statements can be tested to determine what causes the anthills to lean towards the west, in this nature reserve?

A. Anthills are made by termites.
B. Anthills lean towards the west to avoid the sun and the rain.
C. Anthills lean in the direction in which the wind blows.
D. The distribution of anthills depends on the direction of the wind.

3. A study was done to see if leaves added to soil have an effect on tomato production. Tomato plants were grown in four large plots. Each plot had the same kind and amount of soil. One tub had 15 kg of rotted leaves mixed in the soil and a second had 10 kg. A third tub had 5 kg and the fourth had no leaves, added. Each tub plot kept in sun and watered the same amount. The number of kilograms of tomatoes produced in each tub was recorded. What is the hypothesis being tested in this experiment?

A. The greater the amount of sunshine the greater the amount of tomatoes produced.
B. The larger the plot, the greater the amount of leaves added.
C. The greater the amount of water added, the faster the leaves rotted in the plots.
D. The greater the amount of leaves added, the greater the amount of tomatoes produced.

4. Saumu is studying food production in bean plants. She measures food production by the amount of starch produced. She notes that she can change the amount of light, the amount of carbon dioxide, and the amount of water that plants receive. What is a testable hypothesis that Saumu could study in this investigation?

A. The more carbon dioxide a bean plant gets the more starch it produces.
B. The more starch a bean plant produces the more light it needs.
C. The more water a bean plant gets the more carbon dioxide it needs.
D. The more light a bean plant receives the more carbon dioxide it will produce.

5. A poultry farmer raises chickens in cages. He noticed that some chickens lay more eggs than others. A friend of him tells him that, the amount of food and water given to chicken, and the weight of chicken, affect the number of eggs they lay. He decided to conduct a study to prove his friend's claim. Which of the following is NOT a suitable hypothesis for the study?

A. More eggs are produced by chickens that receive more hours of light
B. The more eggs produced by chickens the more weight they loss
C. The larger the cage for chickens the more eggs they will produce
D. The more protein there is in the feed the more eggs produced

6. Doctors noticed that if certain bacteria were injected into a mouse, it developed certain symptoms and died. When the cells of the mouse were examined under the microscope, it was seen that the bacteria did not spread through the body of the mouse, but remained at the area of infection. It was therefore thought that the death is not caused by the bacteria but by certain toxic chemicals produced by them. Which of the statements below provides a possible hypothesis for the cause of death of the mouse?

A. The mouse was killed by the cells that were removed from it to be examined under the microscope.
B. Bacteria did not spread through the body of the mouse but remained at the site of infection.
C. The toxic chemical produced by the bacteria killed the mouse.
D. The mouse was killed by developing certain symptoms.

7. Ibrahim notices that his shower is covered in a strange green slime. His friend Fadhili tells him that coconut juice will get rid of the green slime. Ibrahim decides to check this out by spraying half of the shower with coconut juice. He sprays the other half of the shower with water. After 3 days of the "treatment" there is no change in the appearance of the green slime on either side of the shower. What would be a valid hypothesis for Ibrahim's experiment?

A. Ibrahim believes that coconut juice removes slime, so it must be wrong.
B. 3 days is enough time to remove slime.
C. Coconut juice removes more slime than water.
D. Ibrahim believes that coconut juice removes slime, so it must be right.

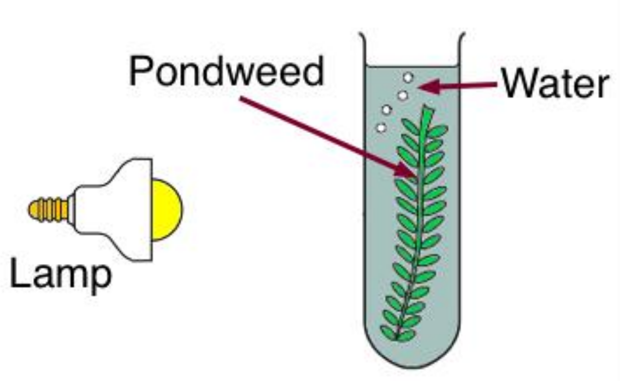
8. Sara wanted to find out if temperature has an effect on the growth of bread mold. She grow the mold in nine containers containing the same amount and type nutrients. Three containers were kept at 0 oC, three were kept at 10 oC and three were kept at room temperature about 27 oC. The containers were examined and the growth of the bread mold was recorded at the end of four days. The dependent variable here is;

A. Amount of nutrients in each container
B. Temperature of the containers
C. The growth of bread mold
D. Number of containers at each temperature

9. One ml of an experimental drug diluted in a saline solution is injected into 20 pregnant mice to determine possible side effects. Which of the following is a suitable "control" for this experiment?

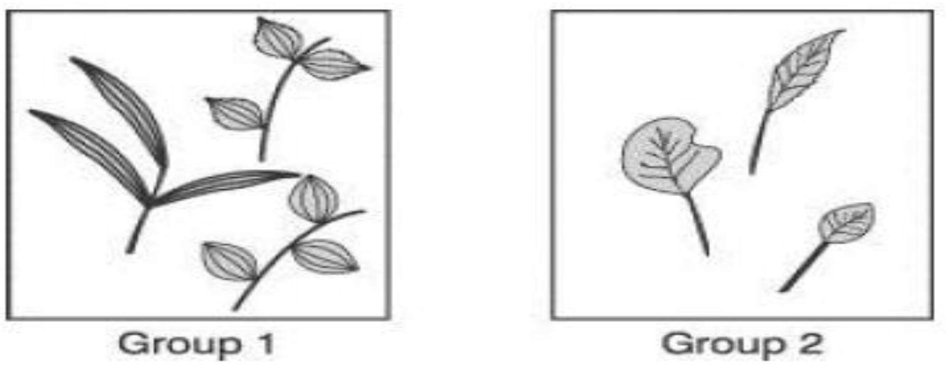
A. 20 male mice injected with 1 ml of the drug
B. 20 pregnant mice injected with 1 ml of saline
C. 20 non-pregnant mice injected with 1 ml of the drug
D. 20 pregnant mice injected with 2 ml of the drug

10. A group of pupils investigated the way in which the colour of light affects photosynthesis.
They put a piece of pondweed into a test tube of water shone light from a lamp with a red light bulb onto the pondweed and then counted the bubbles of gas produced by the pondweed every minute for three minutes. They repeated their experiment using a blue light bulb, a green light bulb and a yellow light bulb. A control is used in order to make the investigation fair. Which one of the following could **NOT** be considered a control?



A. Don't change the distance from the bulb to the test tube.
B. Repeat the experiment at the same time each day for four days.
C. Use the same pondweed for each bulb.
D. Use similar lamps with the same power output.

11. High school students in Biology were conducting a practical experiment on identification of various plants using their leaves as shown in the pictures below. Which two characteristic variables do you think were used to sort the leaves into two groups?



A. shape of leaves and length of stem
B. size of leaves and number of leaves on each stem
C. pattern of veins and number of leaves on each stem
D. pattern of veins and shape of leaves

12. A study was done to see if leaves added to soil have an effect on cabbage production. Cabbage plants were grown in four large tubs. Each tub had the same kind and amount of soil. One tub had 15 kg of rotted leaves mixed in the soil and a second had 10 kg. A third tub had 5 kg and the fourth had no leaves, added. Each tub was kept in sun and watered the same amount. The number of kilograms of cabbage produced in each tub was recorded. What is a controlled variable in this study?

A. Amount of cabbage produced in each tub.
B. Amount of leaves added to the tubs.
C. Number of tubs receiving rotted leaves
D. Amount and kind of soil in each tub.

13. Peter is worried about how the cold winter will affect the growth of his tomatoes. He decided to investigate the effect of temperature on the growth rate of tomato plants. He planted tomato seedlings in four identical pots with the same type of soil and the same amount of water. The pots were put in different glass boxes with different temperatures: One at 0oC, the other at 10oC, and another at room temperature and the fourth at 50oC. The growth rates of the tomato plants were recorded at the end of 14 days. Which of the following factors should be kept constant in this investigation?

A. The time and growth rate of tomato
B. The growth rate of tomato plants and the amount of water used.
C. The type of soil and the amount of water used.
D. The temperature and type of soil used plant

14. Biology students performed an experiment to determine the effect of temperature on heart rate in the crustacean Daphnia. Each group of students exposed Daphnia to varying temperatures from 0 °C to 30 °C and measured the number of heartbeats per minute for three Daphnia after 1 minute of exposure to each temperature. An average was taken for the three Daphnia at each temperature. Identify the independent variable for this experiment.

A. Heart rate
B. Temperature
C. Number of crustaceans
D. Length of exposure

15.Sabrina, a medical scientist is designing an experiment to test the results of a new drug that she hypothesizes will greatly reduce and possibly eliminate the side effects of a new cancer treatment. If this experiment is to be set up correctly, she must

A. Divide the patients into two groups and give each group the same amount of the new drug.
B. Divide the patients into two groups and give one group the new drug and the other group a drug that has no effect (for example, a tablet that only contains sugar).
C. Divide the patients into two groups and give one group the new drug for one week and the other group a different drug for one week.
D. Divide the patients into two groups and give one group the new drug and give the other group nothing.

16. A Biology student is planning a science fair project that involves comparing the effects of two different fertilizers on grass growth. Two 10m² plots are laid out in unshaded lawn and fertilizer A is applied to one plot and fertilizer B to the other plot. The plots are watered equally at regular intervals. After three weeks the grass is mowed and clippings weighted. Which of the following changes in the design would most improve the experiment results?

A. Applying broad spectrum insecticide to both plots
B. Increasing the dimension of the plots from 10m² and hence more grass
C. Adding a plot to which no fertilizer is applied
D. Using plant numbers to determine the differences between fertilizer in Plot A and that in B instead of biomass

17. A gardener is using an electric lamp that is designed to provide enough light for cabbage to grow in a greenhouse. What should the gardener compare the growth of his cabbage with in order to determine how well the lamp works?

A. A similar type of cabbage grown under sunlight
B. A different type of plant grown under sunlight
C. A similar type cabbage grown under the different kind of lamp
D. A different type of plant grown under the same kind of lamp

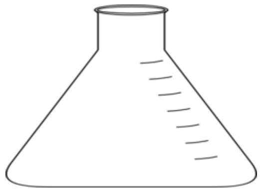
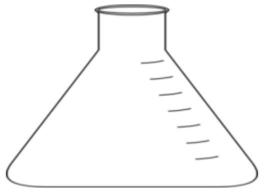
18. A lady grows roses as a hobby. She has six red rose plants and six white rose plants. A friend told her that rose plants produce more flowers when they receive morning sunlight. She reasoned that when rose plants receive morning sunlight instead of afternoon sunlight, they produce more flowers. Which plan should she choose to test her friend's idea?

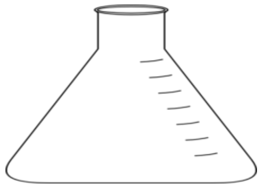
A. Set all her rose plants in the morning sun. Count the number of roses produced by each plant. Do this for a period of four months. Then find the average number of roses produced by each kind of rose plant.
B. Set three red and three white rose plants in the morning sunlight, and three red and three white rose plants in the afternoon sunlight. Count the number of rose flowers produced by each rose plant for four months.
C. Set all her rose plants in the morning sunlight for four months. Count the number of flowers produced during this time. Then set all the rose plants in the afternoon sunlight for four months. Count the number of flowers produced during this time.
D. Set three white rose plants in the morning sunlight and the other three white rose plants in the afternoon sun. Count the number of flowers produced by each white rose plant for four months.

19. A student set up a potometer in the laboratory and measured the rate of movement of water in the capillary. An average of four readings gave a rate of 50mm per minute. The apparatus was then taken outside, where there was a light breeze. Four more readings were taken without delay. The average of these readings was 130 mm per minute. The student concluded that exposure of the shoot to rapid air movement had increased the rate of transpiration. Criticise the design of the experiment?

A. The student conducted an experiment inside and outside the Lab without a testable hypothesis
B. The student ignored the fact that meteorological conditions inside Lab would be different compared to outside.
C. The student conducted a study without having a controlled experimental
D. Apparatus was supposed be taken inside where there was a light breeze

20. Study the description of experimental flasks below and answer the question. A biologist prepares an in vitro analysis of the activity of the enzyme amylase, which promotes the hydrolysis of polysaccharides to monosaccharide residues. Three flasks containing 5millitres of 4 percent amylose (starch) in water are prepared with the addition at time zero of each of the substance as indicated in the flasks below,

i.  ii.  iii.

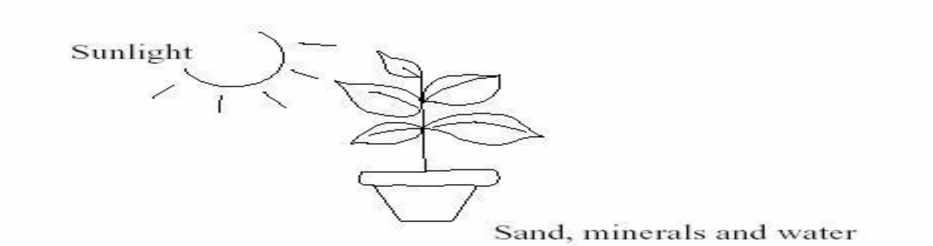


4 % Amylose and 0.5 mil 1%Amylase Solution added	4 % Amylose and 0.5 mil Distilled water added	4 % Amylose and 0.5 mil 1%Amylase Solution added and boiled foe 2 min

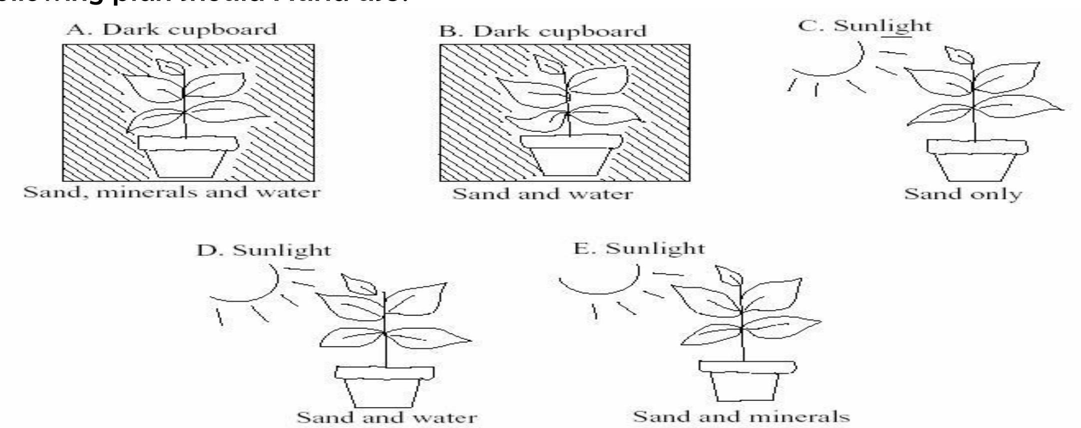
In this experiment to test the effect of amylase on starch, the control would be

A. **flasks A and C**
B. **flask B only**
C. **flask C only**
D. **flask A only**

21. Nuria had an idea that plants needed minerals from the soil for healthy growth. She placed a plant in the Sun, as shown in the diagram below.



In order to check her idea she also needed to use another plant. Which of the following plan should Nuria use?



Nuria should use plan,

A. Plan A
B. Plan B
C. Plan C and E
D. Plan D

22. As part of your research project, you travel to an island to learn more about the habitats and relationships of spiders, centipedes and insects. You and your assistant plotted out five different areas of the island and counted the numbers of spiders, centipedes, and insects living in each plot. Here are your results,

Plot	Spiders	Insects	Centipeders
1	300	25	4
2	426	17	10
3	147	15	21
4	739	78	0
5	79	13	93

From the data, one conclusion that you as a researcher could come up with includes

A. The number of centipedes influences herbivorous insects and spider numbers.
B. Herbivorous insects prefer islands where spiders and centipedes live.
C. Herbivorous insects are not particular about where they live.
D. Spiders are effective at avoiding herbivorous insects.

23. A student placed 20 tobacco seeds of the same species on moist paper towels in each of two petri dishes. Dish A was wrapped completely in an opaque cover to exclude all light. Dish B was not wrapped. The dishes were placed equidistant from a light source set to a cycle of 14 hours of light and 10 hours of dark. All other conditions were the same for both dishes. The dishes were examined after 7 days and the opaque cover was permanently removed from dish A. Both dishes were retuned to the light and examined again at 14 days.

The following data were obtained.

	Dish A		Dish B	
	Day 7 covered	Day14 Uncovered	Day 7 covered	Day14 Uncovered
Germinated seeds	12	20	20	20
Green leaved seedlings	0	14	15	15
Yellow leaved seedlings	12	6	5	5
Mean stem length below first set of leaves	8mm	9mm	3mm	3mm

According to the results of this experiment, germination of tobacco seeds during the first week is,

A. Increased by exposure to light
B. Unaffected by light intensity
C. Prevented by paper towels
D. Accelerated in green-leaved seedlings

24. An experiment to measure the rate of respiration in crickets and mice at 10°C and 25°C was performed using a respirometer, an apparatus that measures changes in gas volume. Respiration was measured in mL of O2 consumed per gram of organism over several five-minute trials, and the following data were obtained.

Organism	Temperature	Average Respiration
Mouse	10	0.0518
Mouse	25	0.0321
Cricket	10	0.0013
Cricket	15	0.0038

According to the data, the mice at 10°C demonstrated greater oxygen consumption per gram of tissue than did the mice at 25°C. This is most likely explained by which of the following statements?

A. A. The mice at 10°C had a higher rate of ATP production than the mice at 25°C.
B. The mice at 25°C weighed less than the mice at 10°C.
C. The mice at 25°C were more active than the mice at 10°C
D. The mice at 10°C had a lower metabolic rate than the mice at 25°C

25. Frogs of three different species are weighed and the amount of oxygen consumed by each species is determined by placing them in a respirometer for 1 hour. The results of this experiment are listed below.

Species	Average Weight in grams	Total Cubic Centimeters of Oxygen consumed in 1 hour
1	15	0.75
2	11	0.55
3	21	1.05

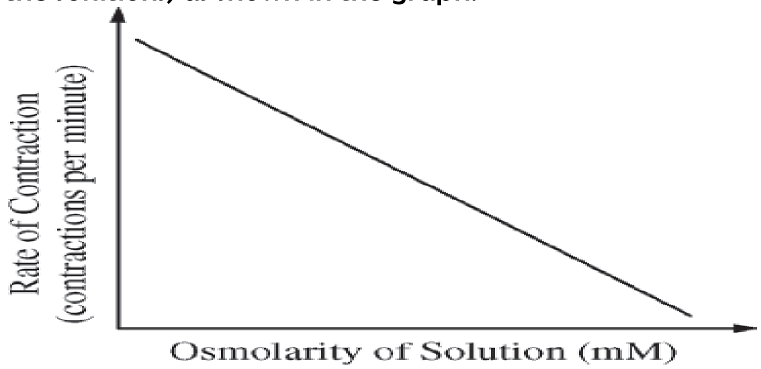
From the information in the table, it is most reasonable to conclude that

A. since all frogs respire through their skin, smaller frogs with smaller surface areas will consume less oxygen per gram of body weight than larger frogs with larger surface areas
B. frogs placed in a warm environment will respire more rapidly than frogs placed in a colder environment
C. each species of frog has its own unique rate of respiration the amount of oxygen consumed per gram of body weight for each species is the same
D. the amount of oxygen consumed per gram of body weight by the largest frog is almost twice that by the smaller frog

26. A student is doing an experiment to determine how change in Acidity affects enzyme activity. The time taken by the disk soaked with Catalase at different acidities to rise to the top of Vial containing 1% Hydrogen peroxide will be measured. If students presents the findings in a line graph which of the following variable will be represented on the x axis of the graph?

A. PH values
B. Catalase concentrations
C. Disk rise times
D. Percentage Hydrogen peroxide

27. Paramecia are unicellular protists that have contractile vacuoles to remove excess intracellular water. In an experimental investigation, paramecia were placed in salt solutions of increasing osmolarity. The rate at which the contractile vacuole contracted to pump out excess water was determined and plotted against osmolarity of the solutions, as shown in the graph.



Which of the following is the correct explanation for the data?

A. At higher osmolarity, lower rates of contraction are required because more salt diffuses into the paramecia.
B. The contractile vacuole is less efficient in solutions of high osmolarity because of the reduced amount of ATP produced from cellular respiration.
C. The contraction rate increases as the osmolarity decreases because the amount of water entering the paramecia by osmosis increases.
D. In an isosmotic salt solution, there is no diffusion of water into or out of the paramecia, so the contraction rate is zero.

28. A Biology teacher wanted to show her class the relationship between light intensity and the rate of plant growth. She carried out an investigation and got the following results.

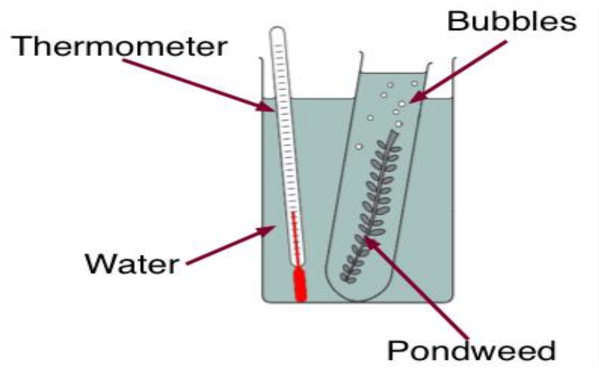
Light intensity(Candela)	Plant growth rate (cm)
250	2
800	5
1000	9
1200	11
1800	12
2000	15
2400	13
2800	10
3100	5

The table above shows the relationship between light intensity and the growth rate of a plant.

Which of the following statements correctly describes what these results show?

A. As light intensity increases, plant growth also increases.
B. As plant growth increases, light intensity decreases.
C. As plant growth increases, light intensity increases then decreases.
D. As light intensity increases, plant growth increases then decreases.

29. A student investigated the effect of temperature on the rate of photosynthesis in pondweed. The diagram shows the way the experiment was set up. What two measurements would the student make to operationally define and calculate the rate of photosynthesis?



A. Volume of water and the number of bubbles.
B. Temperature of the water and the number of bubbles.
C. Volume of water and the temperature of the water.
D. Temperature of the water and the number of leaves

30. In a Floriculture farm, a gardener notices that her Chrysanthemums plants are being attacked by unknown kind of insect. She needs to get rid of the insect. An extension agent in that area told her that "Insect-Away" powder is the best insecticide to use. Her brother says "Chrysanthemums-Saver" spray might work best in that situation. The gardener decided to select six Chrysanthemums plants and applies the powder to three and the spray to three plants. A week later she counts the number of live insects on each of the plants. How is the effectiveness of the insecticides measured in this study?

A. Counting the number of insects remaining on the plants.
B. Measuring the amount of spray or powder used.
C. Determining the condition of the plants after spraying or dusting.
D. Weighing the Chrysanthemums each plant produces.

31. After a Respiration class, a Form V Biology teacher wanted to guide her students find out the effect of exercise on pulse rate. He divided his students into four groups. He asked each of three groups of learners to do some push-ups over a given period of time, and then measure their pulse rates: one group did the push-ups for two minutes; the second group for four minutes; the third group for six minutes and then a fourth group did not do any push-ups at all. How can students and their teacher best measure pulse rate in this investigation?

A. By counting the number of push-ups in one minute.
B. By counting the number of pulses in one minute.
C. By counting the number of push-ups done by each group.
D. By counting the number of pulses per group.

32. Margret wants to speed up the production of tomato plants to meet the growing market demand. She plants tomato seeds in several trays. Her hypothesis is that the more moisture seeds receive the faster they sprout. How can she test operationally her hypothesis?

A. Count the number of days it takes seeds receiving different amounts of water to sprout.
B. Measure the height of the tomato plants a day after each watering.

C. Measure the amount of water used by plants in different trays.
D. Count the number of tomato seeds placed in each of the trays.

33. Sofia wants to find out the amount of water contained in meat, cucumber, cabbage and maize grains. She finely chopped each of the foods and carefully measured 10 grams of each. She then put each food in a dish and left all the dishes in an oven set at 100oC. After every 30 minutes interval, she measured the mass of each food, until the mass of the food did not change in two consecutive measurements. She then determined the amount of water contained in each of the foods. How is the amount of water contained in each food measured in this experiment?

A. By finding the difference between the original and the final mass of each food.
B. By heating the samples at a temperature of 100oC and evaporating the water.
C. By measuring the mass of the foods every 30 minutes and determining the final mass.
D. By finely chopping each food and measuring 10 grams of it, at the beginning of the investigation.

34. A story in Mwananchi newspaper reports that post-surgical patients who received a type of alternative therapy involving light touching of the body reports less severe pain than those who did not receive the therapy. The writer concludes that this therapy should be adopted in all hospitals. Which information would be most useful to the reader in assessing the validity of this conclusion?

A. The percentage of patients in all hospitals who reports that they have experienced severe pain after surgery
B. The estimated cost to hospitals of providing alternative therapy to all of their postsurgical patients
C. The percentage of patients reporting reduced pain after receiving other forms of attention such as traditional massage or talking with doctors
D. The level of pain reported by postsurgical patients before and after they received the alternative therapy

35. A Biology student tests this hypothesis: the greater the amount of vitamins given to rats the faster they will grow. How can this student measure how fast rats will grow?

A. Weigh the rats every day.
B. Measure the speed of the rats.
C. Measure the amount of exercise the rats receive.
D. Weigh the amount of vitamins the rats will eat.

THE END

BIOLOGY PROCESS SKILLS TEST (BPST)

ANSWER GRID

STUDENTS NUMBER..........................SCHOOL NAME............................... Sex ....

Question Number	Correct Answer Option	Question Number	Correct Answer Option
1		23	
2		24	
3		25	
4		26	
5		27	
6		28	
7		29	
8		30	
9		31	
10		32	
11		33	
12		34	
13		35	
14			
15			
16			
17			
18			
19			
20			
21			
22			

Appendix II: Integrated Science Process Skills Test (TIPS-II)

INTEGRATED SCIENCE PROCESS SKILLS TEST II (TIPS II)

ANSWER GRID

STUDENTS NUMBER……………………SCHOOL NAME……………………………………

GRADE LEVEL ……………………………SEX…………………………………………………

Question Number	Correct Answer Option	Question Number	Correct Answer Option
1		23	
2		24	
3		25	
4		26	
5		27	
6		28	
7		29	
8		30	
9		31	
10		32	
11		33	
12		34	
13		35	
14		36	
15			
16			
17			
18			
19			
20			
21			
22			

Appendix III: Content Validation Form

Test Content Validation Form

Judge: ___________________

Title: ___________________

Date: 13 July 2014

This is an evaluation form to get information of how valid is this test developed to assess the level of Biology students' science process skills. Please read the prepared guidelines for item objectivity, accuracy and clarity determine whether the items represent the intended objective and construct. For each test item, please make judgments and score on a scale of 0 (low) to 3 (high) to indicate your opinion. The following criteria are item assessment guidelines to be used.

For item objectivity (relevance)

i. Is the item giving a description of an investigation for students to identify the dependent and independent variables? (Identifying and controlling variables)
ii. Is the item giving description of an investigation for students to identify how variables are operationally defined? (Operational definitions)
iii. Is the item giving a problem with a dependent and independent variable specified for students to identify the variables which may affect it? (Identifying and controlling variables)
iv. Is the item giving a problem with dependent variables and a list of possible independent variables for students to identify testable hypothesis? (Stating hypothesis)
v. Is the item giving a verbally described variable for students to select a suitable operational definition for it? (Operational definitions)
vi. Is the item giving a problem with a dependent variable specified for students to identify a testable hypothesis? (Stating hypothesis)
vii. Is the item giving a hypothesis for students to select a suitable design for an investigation to test it? (Designing investigations)
viii. Is the item giving a description of an investigation and obtained result or data for students to identify a graph that represent the data? (Graphing and interpreting data)
ix. Is the item giving a graph or table of data from an investigation for students to identify the relationship between variables? (Graphing and interpreting data)
x. Is the item based on instructional objectives as specified by the Tanzania Biology Curriculum?
xi. Is the item emphasizing higher level thinking?

For item accuracy (clarity)

i. Is the items focused on a single problem?
ii. Are the alternatives free from clues as to which response is correct?
iii. Items are kept independent of one another and hence no item cuing?
iv. Is the central idea and most of the phrasing included in the stem and not in the options?
v. Are the distracters enough and plausible?
vi. Are options placed in logical or numerical order?
vii. Is there is only one correct answer among alternatives?
viii. Are distracters are plausible and functional?
ix. Is the use of specific determiners such as *never* and *always* avoided?
x. Are the alternatives similar in length?
xi. Is the wording of alternatives clearly and concisely?

For item ambiguity

i. Is the vocabulary is kept consistent with the group of students being tested (secondary students) ?
ii. Is the excess verbiage in the stem avoided?
iii. Is the grammar in the alternatives consistent with the stem?
iv. Is the used sentence structure simple for students?
v. How would you rate the use proper grammar, punctuation, and spelling in the item?
vi. Is the use of unnecessarily difficult vocabulary avoided in the item?
vii. Is the lay out the item in a clear and consistent manner?
viii. Is the wording lets the examinee know exactly what is being asked?
ix. Is the direction for the stem clear?

Test Content Validation Form

Judge: ______________________ Title: ______________________ Date: ______________

	OBJECTIVITY						ACCURACY					CLARITY				
	0.0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0	2.2	2.4	2.6	2.8	3.0
1																
2																
2																
3																
4																
5																
6																
7																
8																
9																
10																
11																
12																
13																
14																
15																
16																
17																
18																
19																
20																
21																
22																
23																
24																
25																
26																
27																
28																
29																
30																
31																
32																
33																
34																
35																

Appendix IV: Genetics Test

GENETICS COGNITIVE TEST

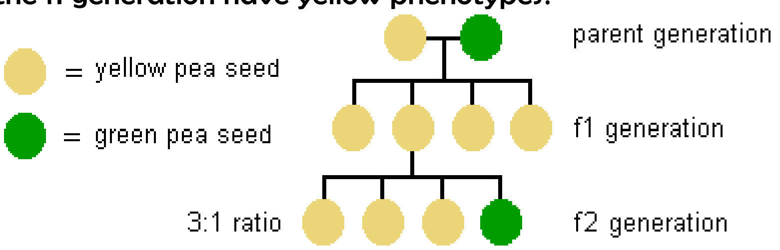
INSTRUCTIONS
DURATION: 40 minutes

iv. DO NOT WRITE ANYTHING IN THIS PAPER
v. ANSWER ALL THE QUESTIONS ON THE ANSWER GRID PROVIDED, BY PUTTING A CORRECT ALTERNATIVE OF YOUR CHOICE.
vi. PLEASE DO NOT GIVE MORE THAN ONE ANSWER PER QUESTION.

1. What is the relationship among DNA, a gene, and a chromosome?
   A. A chromosome contains hundreds of genes which are composed of DNA.
   B. A gene contains hundreds of chromosomes which are composed of protein.
   C. A chromosome contains hundreds of genes which are composed of protein.
   D. A gene contains hundreds of chromosomes which are composed of DNA.

2. In DNA, a single strand of deoxyribonucleotides are held together by __________ bonds, and the two strands of deoxyribonucleotides are held together by __________ bonds between adjacent N-bases.
   A. covalent; peptide
   B. covalent; hydrogen
   C. peptide; hydrogen
   D. hydrogen; peptide

3. Assuming that both parent plants in the diagram below are homozygous, why would all of the f1 generation have yellow phenotypes?



   A. Because the f1 genotypes are homozygous
   B. Because yellow is dominant over green
   C. Because both parents passed on yellow alleles
   D. Because of the interaction of heredity and environment

4. A cross between two true breeding lines one with dark blue flowers and one with bright white flowers produces F1 offspring that are light blue. When the F1 progeny are selfed a 1:2:1 ratio of dark blue to light blue to white flowers is observed. What genetic phenomenon is consistent with these results?
   A. epistasis
   B. incomplete dominance
   C. codominance
   D. random mating

5. Why are there more males with color blindness than females?
   A. The gene for color blindness is found on the Y chromosome.
   A. The recessive gene is usually masked by another X chromosome in females.
   B. Color blindness is an X-linked dominant trait.
   C. All the sons of an affected male will have the disorder.

6. Musa and Neema are planning a family, but since each has a brother who has sickle-cell anemia, they are concerned that their children may develop sickle-cell disease. Neither Musa nor Neema and their respective parents have the disease. They consult a genetic counselor who probably told them,

A. That all of their children will have sickle-cell disease
B. There is very little chance that any of their children will have sickle-cell disease
C. That one out of four of their children could be expected to have sickle cell-disease
D. That it is possible that none of their children will have the disease but blood tests on them both will be required to make sure

7. In people with sickle cell disease the red blood cells breakdown, clump, and clog the blood vessels. The broken cells accumulate in the spleen. Among other things this leads to physical weakness, heart failure, pain, brain damage and spleen damage. Affected individuals become paralyzed and can develop rheumatism, pneumonia and other diseases and kidney failure. This is an example of

A. the polygenic nature of sickle cell disease
B. the pleiotropic effects of the sickle cell allele
C. an epistatic interaction between the sickle cell allele and a proteolytic enzyme gene
D. infectious organisms acting on the sickle cell allele

8. You set up an experiment in which you breed two populations of true-breeding pea plants. The first true-breeding population has yellow round seeds and the second has green wrinkled seeds. All of the F1 plants yield yellow round seeds. When you self fertilize the F1 the F2 generation yields a mixture of yellow round, yellow wrinkled, green round and green wrinkled seeds. What does this tell you about the alleles for seed color and shape?

A. the recessive alleles are always expressed
B. the alleles are on different chromosomes
C. the two alleles for each character segregate during gamete production
D. both genes are on the same chromosome

9. If the parents are AO and BO genotypes for the ABO blood group, their children could include which of the following genotypes?

A. AO, BO, and AB only
B. AA, BB, and AB only
C. AO, BO, and OO only
D. AO, BO, AB, and OO only

10. While on a field trip in the jungle you find a new species of mouse. You catch a pair and take them back to the lab. In mice, black coat color, B, is dominant to brown b, yet the female mouse gives rise to a large litter in which 9 of the offspring were black, 3 were brown and 4 were white. You conclude that

A. a new mutation has occurred in the mice
B. this is an example of polygenic inheritance
C. there must be an epistatic interaction influencing coat color
D. the coat color alleles are codominant

11. In a particular species of mammal black hair (B) is dominant to green hair (b) and red eyes (R) are dominant to white eyes (r). If a BbRr individual is mated with a bbrr individual the expected phenotypic ratio of the offspring is 1 black-red : 1 black-white: 1 green-red : 1 green-white. However, when you mate these individuals you find that the phenotypic ratio of the offspring is 6 black-red : 1 black-white : 1 green-red : 6 green-white. What could account for this difference?

A. The genes for hair color and the genes for eye color are carried on different chromosomes
B. The expected results did not take genetic recombination into account
C. The genes for hair color and eye color are linked
D. The genes for hair color and eye color show dependent assortment

12. Haiti is settled by peoples of both African and European ancestry. A young couple, both with mixed ancestry, marry and have several children. The children vary widely in the amount of skin melanin production, with one child being lighter than either parent, and one being darker. The simple explanation for this is

A. multiple alleles are available for the one chromosomal locus that governs skin color.
B. the environment affected the phenotype that developed.
C. *polygenic inheritance.*
D. gene linkage.

13. Humans have 23 pairs of chromosomes, while our closest relatives, chimpanzees, have 24. Chromosome studies indicate that at some point early in human evolution, two chromosomes simultaneously broke into a large portion and a small portion. The large parts combined to form a large chromosome, and the small parts combined to form a much smaller chromosome (which was subsequently lost). This important chromosomal change could best be described as

A. nondisjunction followed by deletion
B. translocation followed by deletion
C. duplication followed by deletion
D. translocation followed by inversion

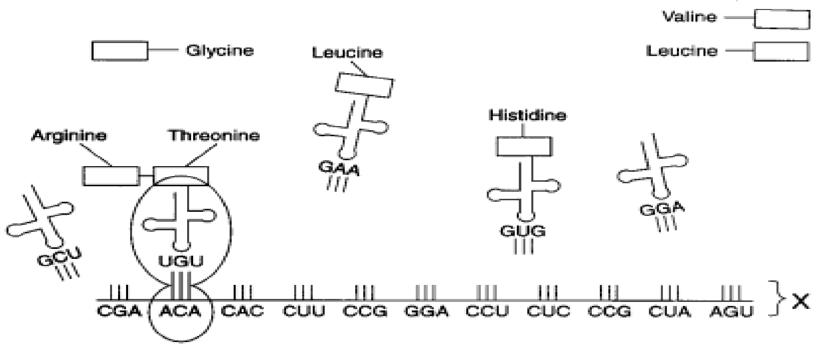
14. A geneticist isolates a gene for a specific trait under study. She also isolates the corresponding mRNA. Upon comparison, the mRNA is found to contain 1,000 fewer bases than the DNA sequence. Did the geneticist isolate the wrong DNA?

A. yes, mRNA is made from a DNA template and should be the same length as the gene sequence
B. yes, the mRNA should contain more bases than the DNA sequence because bases flanking the gene are also transcribed
C. no, the final mRNA contains only exons, the introns were removed
D. no, the mRNA was partially degraded after it was transcribed

15. Haemophilia A is a severe coagulation disorder that shows X-linked recessive inheritance. Red-green colour blindness also shows X-linked recessive inheritance. A man with both haemophilia A and colour blindness is referred for genetic counseling. Assume that his partner is not a carrier of either of these conditions. Which of the following is correct?

A. The probability that each of his daughters will be a carrier of haemophilia A is 1 in 2
B. The probability that each of his daughters will be a carrier of colour blindness is 1 in 2
C. The probability that each of his daughters will be a carrier of both conditions is 1
D. The probability that each of his sons will be affected with haemophilia A is 1 in 2

Use the diagram below for Questions 16-17



16. Structure X was made in the

A. nucleus
B. cytoplasm
C. lysosome
D. vacuole

17. Which amino acid would be transferred to the position of codon CAC?

A. leucine
B. glycine
C. valine
D. histdine

18. What is the role of tRNA during translation?

A. bond to open the DNA strand to carry the code for protein synthesis out of the nucleus
B. carry ribosomes to the site of protein synthesis
C. break aparty mRNA and send it back to the nucleus so that it can be reused
D. Carry amino acids to the mRNA for correct placement into the protein chain

19. Some plants fail to produce chlorophyll, and this trait appears to be recessive. If we locate a plant that is heterozygous for this trait, self-pollinate it and harvest seeds, what are the likely phenotypes of these seeds when they germinate?

A. All will be green with chlorophyll since that is the dominant trait.
B. All will be white and lack chlorophyll since this is self-pollinated.
C. About one-half will be green and one-half white since that is the distribution of the genes in the parents.
D. About one-fourth will be white and three-fourths green since it is similar to a monohybrid cross.

20. If you had two guinea pigs of opposite sex, both homozygous, one black and one brown, but you didn't know which was the dominant characteristic, how would you find out the dominant color?

A. *Mate them together and see what color the offspring are--that will be the dominant color.*
B. Mate them together and see what color the offspring are--the other will be the dominant color.
C. Mate them together , then mate their offspring to see what color the grandchildren are--that will be the dominant color.
D. Mate them together, then mate their offspring to see what color the grandchildren are--the other color will be the dominant color.

21. In 1940, two researchers named Weiner and Landsteiner discovered that about 85 percent of the human population sampled possessed a blood cell protein that had been previously detected in Rhesus monkeys. This blood type was labeled Rh positive, and Rh^+ was found to be dominant over the absence of the blood factor (Rh). Under normal Mendelian inheritance, which of the following statements is FALSE?

A. Two Rh^+ parents could have an Rh^- child.
B. Two Rh^- parents could have an Rh^+ child.
C. An Rh^- child would require that both parents be carriers of at least one Rh^- gene.
D. It is possible with just one pair of parents to have children where some siblings are Rh^- and some are Rh^+.

22. Since each child of two heterozygous parents has a 1/4 chance of receiving a recessive trait from each parent,

A. if the first child is phenotypically recessive, then the next child must be phenotypically dominant.
B. if the first child is phenotypically recessive, then the next child has a 3/4 chance of being phenotypically recessive.
C. if the first child is phenotypically recessive, then the next child has a 1/2 chance of being phenotypically recessive.
D. *no matter what the first child's phenotype, the next child will have a 1/4 chance of being phenotypically recessive.*

23. In pea plants, the gene for round seed (R) is dominant, and wrinkled seeds (r) are recessive. The endosperm of the pea is also either starchy, a dominant gene (S), or waxy (s). What can be said of a fully heterozygous (or dihybrid) cross?

A. It is impossible to secure offspring that are homozygous for both dominant genes.
B. It is impossible to secure offspring that are homozygous for both recessive genes.
C. It is impossible to secure offspring that are homozygous for one dominant gene such as round seed and homozygous recessive for the other recessive waxy gene.
D. All of these choices are possible combinations in a dihybrid cross.

24. If a human who is a tongue roller (T) and has unattached ear lobes (E) marries a person who cannot roll their tongue and has attached earlobes, could they produce an offspring that was also a non-tongue roller with attached earlobes? What would be the genotype of the first parent? the second parent?

A. yes; TtEe; ttee
B. yes; TtEE; ttEe
C. no; TTEE; ttee
D. yes: TTEe; ttee

25. Lethal genes (genes that result in the failure to develop a vital organ or metabolic pathway) are nearly always recessive. Animal breeders who discover a unique trait and cross-breed to increase the occurrence of that trait often encounter a noticeable increase in lethal genes. Why?

A. The lethal recessive gene may be linked to the desire trait gene.
B. Spreading the gene among offspring of both sexes will increase the likelihood it will be sex-linked and expressed.
C. *The cross-mating of closely related individuals, or inbreeding, increases chances the two recessive genes will "meet" in offspring.*
D. "Pleiotropy"- the gene that is being selected for may have the second effect of being lethal.

GENETICS TEST

ANSWER GRID

STUDENTS NUMBER..........................SCHOOL NAME..

GRADE LEVEL.................................... SEX...

Question Number	Correct Answer Option	Question Number	Correct Answer Option
1		23	
2		24	
3		25	
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			

Appendix V: FSWEx Self-concept scale by Kasrten (2012)

FSWEx (Damerau 2012)

This questionnaire contains statements about your willingness in participating in this experimental (science process skills) class. You will be asked to express your agreement on each statement. There are no "right " or "wrong" answers. Your opinion is what is wanted. Think about how well each statement describes your willingness in this experiment(science process skills) .Tick either strong disagree, disagree, no opinion, agree and strong agree. Be sure to give an answer for all questions. If you change your mind about an answer, just cross it out and tick another. Simply give your opinion about all statements.

Your Number___________________; School_____________________________________; Form_________
Male_______________ Female_________

	A. ***planning experiments***	**Strongly disagree**	**disag ree**	**some times**	**agree**	**Strongly agree**
i	In my daily life it often happens that questions emerge which can be solved by experiments					
ii	It is easy for me to formulate theoretically based hypotheses.					
iii	It is quite easy for me to develop an experiment to solve a given problem.					
iv	I´m good at choosing suitable laboratory equipment for experiments.					
v	It is easy for me to develop an experimental instruction to solve a specific scientific research question.					
vi	I find it easy to transfer an idea for an experiment into an experimental setting.					
	B. carrying out experiments					
i	I don´t have a good hand for carrying out experiments					
ii	I am good at working with laboratory equipment.					
iii	Writing down experimental observation is always hard for me.					
iv	I have no problem with arranging experimental setups.					
v	I am good at doing experiments.					
vi	Handling lab equipment is very easy for me.					
	Analzsing data					
i	Analysing experimental data is easy for me.					
ii	I often have problems with interpreting experimental results.					
iii	Interpreting experimentel observation is easy for me.					
iv	I do well in analysing experimental results.					
v	Detecting possible errors in an experiment that went wrong is easy for me.					
vi	I can easily generate graphs based on experimental data.					

1

Appendix VI: Modified Science Motivation Questionnaire (SMQ-II) by Glynn, et al. (2011)

SCIENCE MOTIVATION QUESTIONNAIRE- II (SMQ-II)

In order to better understand what you think and how you feel about your science courses, please respond to each of the following statements from the perspective of "When I am in a science eocess skills course..."

Your Number__________________; School______________________________________; Form__________
Male_______________ Female_________

	Statements	Never	Rarely	Sometimes	Often	Always
1	The science process skills I learn is relevant to my life.					
2	I like to do better than other students on science process skills tests.					
3	Learning science process skills is interesting.					
4	Getting a good science process skills grade is important to me.					
5	I put enough effort into learning science process skills.					
6	I use strategies to learn science process skills well.					
7	Learning science process skills will help me get a good job.					
8	It is important that I get an "A" in science process test.					
9	I am confident I will do well on science process skills tests.					
10	Knowing science process skills will give me a career advantage.					
11	I spend a lot of time learning science process skills.					
12	Learning science process skills makes my life more meaningful.					
13	Understanding processes of science will benefit me in my career.					
14	I am confident I will do well on science labs and projects.					
15	I believe I can master science knowledge and skills.					
16	I prepare well for science process skills tests and labs.					
17	I am curious about discoveries in science process skills.					
18	I believe I can earn a grade of "A" in science process skills.					
19	I enjoy learning science process skills.					
20	I think about the grade I will get in science process skills.					
21	I am sure I can understand science process skills.					
22	I study hard to learn science process skills.					
23	My career will involve science process skills.					
24	Scoring high on science process skills tests and labs matters to me					
25	I will use science problem-solving skills in my career.					

Appendix VII: Supervisor´s Research Permission

BERGISCHE UNIVERSITÄT WUPPERTAL

Prof´in Dr. Angelika Preisfeld
Zoologie und Biologiedidaktik
Fachbereich C – Biologie

Gaußstraße 20
42119 Wuppertal

Bergische Universität Wuppertal, Prof´in Dr. A. Preisfeld, Gaussstr. 20, 42119 Wuppertal

RAUM	L.10.09
TELEFON	+49 (0)202 439 2815
FAX	+49 (0)202 439 2967
MAIL	apreis@uni-wuppertal.de
WWW	www.biologie.uni-wuppertal.de
SEKRETARIAT	Frau Marion Litz 13-16:45 Uhr
TELEFON	+49 (0)202 439 2974
DATUM	Wuppertal, 21. Mai 2014

Doctoral candidate research clearance for Jamal Athuman

Mr. Jamal Athuman (Morogoro) is doing research as a PhD student at the University of Wuppertal, Germany supported by a MoEVT/DAAD scholarship. He is also an employee of the Sokoine University of Agriculture.

In the course of his research he needs to conduct research at several schools. I hereby ask you politely to grant him any support he needs to carry out his research successfully in the time from 1.07.2014 to 31.03.2017 in the Morogoro region.

In case further information is needed, feel free to contact me or the Office of the Sokoine University:

Office of the Vice Chancellor:
P.O. Box 3000, Chuo Kikuu, Morogoro- Tanzania
Tel. + 255 23 2604651 & + 255 23 2604523
Fax: + 255 23 2604651
E-mail: vc@suanet.ac.tz

With kind regards,

Prof. Dr. Gela Preisfeld

Appendix VIII: Research Permission from Morogoro Regional Administration

THE UNITED REPUBLIC OF TANZANIA

PRIME MINISTER'S OFFICE
REGIONAL ADMINISTRATION AND LOCAL GOVERNMENT

Telegraphic Address: "REGCOM"
Phones: 023 2604237/2604227

Fax No: 260 09 73
In Reply please quote:

Regional Commissioner's Office,
P.O. Box 650,
MOROGORO.

Ref. No: AB.175/245/01/224

09 July, 2014.

District Administrative Secretaries,
Morogoro and Mvomero.

RE: RESEARCH PERMIT

Please refer to the above captioned subject.

I have a great honour to introduce to you **Mr Jamal Athuman** a PhD student at the University of Wuppertal, Germany who at the moment conduct a research in our Region.

The title of the research in question is **"Development and Validation of Science Process Skills test for Measuring students ability and the effectiveness of Inquiry Based Approach. A case of Morogoro and some Experience from NRW Germany".**

The permit is granted from July, 2014 to March, 2017 will cover **Morogoro and Mvomero Districts.**

Please provide him with all needed assistance to enable the accomplishment of this research.

Thank you for your cooperation.

Tumaini Wapalila
For: REGIONAL ADMINISTRATIVE SECRETARY
MOROGORO

Copy:√Researcher

Appendix IX: Research Permission from Morogoro municipality

JAMHURI YA MUUNGANO WA TANZANIA
OFISI YA WAZIRI MKUU
TAWALA ZA MIKOA NA SERIKALI ZA MITAA

Anuani ya simu:" **MKUU WA WILAYA**
Simu Nambari: **2614096**
Fax Nambari: **2613848**

Unapojibu tafadhali taja:

Kumb. Na. AB.210/249/01/139

OFISI YA MKUU WA WILAYA
S.L.P 681
MOROGORO

10 Julai, 2014

Mkurugenzi,
Halmashauri ya Manispaa,
S.L.P 166,
MOROGORO.

YAH: UTAFITI

Somo tajwa hapo juu lahusika.

Napenda kumtambulisha kwako **Bwana Jamal Athuman** ambaye ni mwanafunzi Chuo Kikuu cha **Wuppertal, Germany**.

Anategemea kufanya utafiti katika mkoa wetu katika shule zote za Sekondari za Halmashauri ya Manispaa Morogoro. Anafanya juu ya **"Development and Validation of Science Process Skills test for Measuring students ability and the effectiveness of Inquiry Based Approach. A case of Morogoro and some Experience from NRW Germany".**

Kibali kinatolewa kuanzia Julai, 2014 hadi March, 2017.

Tafadhali apewe ushirikiano wa kutosha ili aweze kukamilisha utafiti huo.

pp
Tamari P. Isike
Kny: KATIBU TAWALA WILAYA
MOROGORO

K.n.y. KATIBU TAWALA WILAYA
MOROGORO

Nakala:
1. Afisa Elimu Sekondari
   HALMASHAURI YA MANISPAA
   MOROGORO
2. Jamal Athuman
   MTAFITI

Appendix X: Sample of inquiry-based lessons for the experimental group students

BIOLOGY – Activity Names ______________________________

Modeling Meiosis ___________________________

Period 1 2 3 4 5 6 7 8 ___________________________

Date: _____________

Name of the school _______________________

INTRODUCTION

The body cells of plants and animals are diploid. A **diploid** (*2n*) cell has two sets of chromosomes in its nucleus. A cell with only one set of chromosomes in its nucleus is termed **haploid** (*n*). Egg and sperm, **gametes**, are examples of haploid cells. When gametes fuse at fertilization, a diploid **zygote** is formed. The zygote contains one set of chromosomes from each parent.

The processes that produces haploid (*n*) cells such as gametes from diploid (*2n*) cells is called **meiosis**. Before meiosis begins, DNA replication occurs. Following replication, each chromosome consists of two **chromatids** that are joined by a **centromere.** Meiosis involves two successive divisions of the nucleus. The first of these divisions is called **meiosis I.** During meiosis I, the **homologous chromosomes** (chromosomes that carry the same genes and are similar in size and shape) come together, or pair up, and then separate. The nuclei that result from meiosis I contain only one set of chromosomes, or one chromosome from each pair of homologous chromosomes. Therefore, meiosis I is also known as reduction division because each of the resulting nuclei contains half the number of chromosomes of the original cell. The second division of the nucleus is called **meiosis II**. During meiosis II, the chromatids separate. forming 4 haploid nuclei.

During meiosis I, the chromatids of a **homologue** (member of a pair of homologous chromosomes) may exchange parts. This exchange of segments between chromatids is called **crossing over**. Crossing over, as well as the fusion of two gametes during sexual reproduction, is a type of **genetic recombination**, which is the regrouping of genes into new combinations.

OBJECTIVES

To model the stages of meiosis in an animal cell
To demonstrate genetic recombination
To relate the events of meiosis to the formation of haploid gametes

MATERIALS

4 pieces of string (1 meter long) scissors paper clips (8)

4 pieces of string (40 cm long) metric ruler tape

8 pieces of string (10 cm long)

4 strips of paper (2cm x 6 cm), one each of light blue, dark blue, light green, dark green

PROCEDURE

1. Using a 1 meter piece of string, make a circle on the lab table to represent the cell membrane of a cell. Using a 40 cm piece of string, make another circle inside the cell to represent the nuclear membrane.

2. Fold each of 4 strips of paper (one light blue, one dark blue, one light green, and one dark green) in half lengthwise. Then place each of these folded strips inside the nucleus to represent the four chromosomes before replication. The light and dark strips of the same color represent homologous chromosomes. The light strips represent chromosomes from one parent and the dark strips, chromosomes from the other parent.

3. Interphase. To represent DNA replication, unfold each paper strip and cut each in half lengthwise. The two pieces that result from cutting each homologous strip represent the chromatids. Attach the two identical chromatid strips at the center with a paper clip so that an X is formed (see Fig. 1 below). Each paper clip represents a centromere.

What process did you model when you cut the paper strips in half?

____

What is the function of the centromere?

____

4. Prophase I. Remove the nuclear membrane (the 40 cm string). Place the blue chromatid pairs next to each other and the green chromatid pairs next to each other. Simulate crossing over by measuring and cutting a 1 cm piece from the tip of a light blue strip and a dark blue strip. Tape the light blue piece to the dark blue strip and the dark blue piece to the light blue strip (see Fig. 1). Repeat this procedure with the green strips.

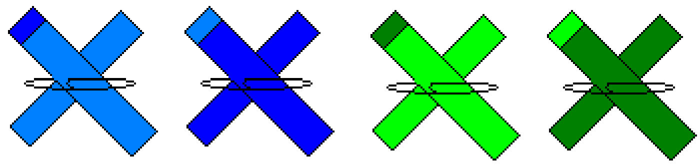


Fig. 1

What is the purpose of placing the light and dark strips of the same color side by side?

____

5. Metaphase I. Place four 10 cm pieces of string inside the cell so that two strings extend from one side into the center of the cell and the other two strings extend from the opposite side into the center of the cell. These strings represent the spindle fibers. Using a small piece of tape, attach one string to the centromere of each of the four chromatid pairs. Move the chromatid pairs to the center of the cell so that they line up in a double file of X's, blue next to blue and below them, green next to green. Make sure that strings attached to similar colors come from opposite sides of the cell.

6. Anaphase I. To simulate anaphase I, gather the loose ends of the two strings on each side of the cell and gently pull the strings in opposite directions so that the homologous pairs of chromosomes are moved to opposite sides of the cell.

7. Telophase I. Carefully remove the tape from each of the centromeres. Place a 40 cm piece of string around each group of chromatids, forming two nuclei. Remove the original 1 meter piece of string and place new 1 meter pieces of string around each of the nuclei thus forming two cells.

How many chromosome pairs are in each of the cells you formed? __________

List the materials used to make these two cells and what each represents.

____

____

8. Prophase II. Remove the strings that represent the nuclear membranes of each cell. Attach a 10 cm piece of string to each <u>chromatid</u> (not the centromere).

What must happen to the centromeres before the chromatids can separate?

____

9. Metaphase II. Move the chromatid pairs to the center of each cell and line them up in a column with the blue X above the green X. Make sure the strings attached to each of the chromatids come from opposite sides of each cell.

10. Anaphase II. Gather the strings on both sides of each cell and gently pull in opposite directions, separating the paper strips (chromatids) and pulling them to opposite sides of each cell. *Note: only one strip in each pair should have the paper clip attached*.

11. Telophase II. Remove all of the strings and the paper clips. Each strip of paper now represents a chromosome. Place a 40 cm piece of string around each of the 4 groups of chromosomes, thus forming 4 nuclei. Place a 1 meter piece of string around each of the nuclei thus forming 4 cells.

How many chromosomes are in each of the cells you formed? Are these cells haploid or diploid? ____

12. Save the paper clips and dispose of all the strings and paper strips you cut. Make sure your work area is returned to the way you found it.

ANALYSIS

1. What is the diploid number of the original cell you modeled? How many homologous pairs does this represent?

____

2. If a cell with a diploid number of 6 undergoes meiosis, what will the cell look like after Telophase I? Draw it in the space below and label all parts.

3. Give two reasons why meiosis is important in sexual reproduction.

____

____

4.Why is meiosis I known as reduction division?

____

____

5. List two ways that meiosis is different from mitosis.

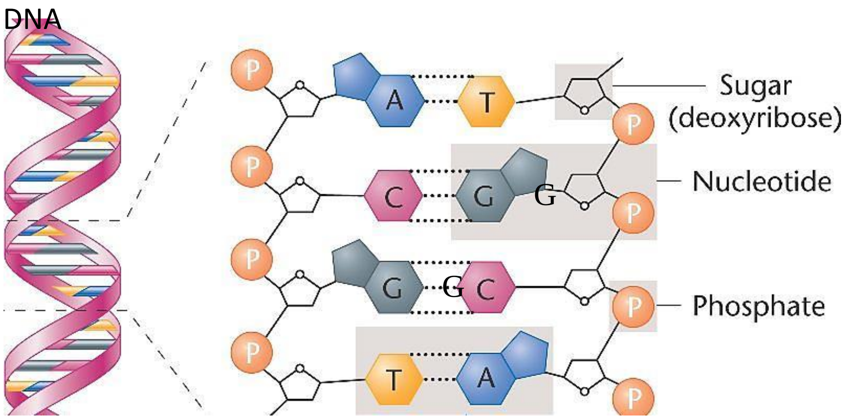
____

____

DNA Structure and Function

This drawing shows a short section of a DNA double helix with a diagram of four nucleotides in each strand of the double helix. Each nucleotide has:

- a phosphate group (P) and a sugar molecule in the backbone of the DNA strand
- one of the four bases (**A** = adenine, **C** = cytosine, **G** = guanine, or **T** = thymine)

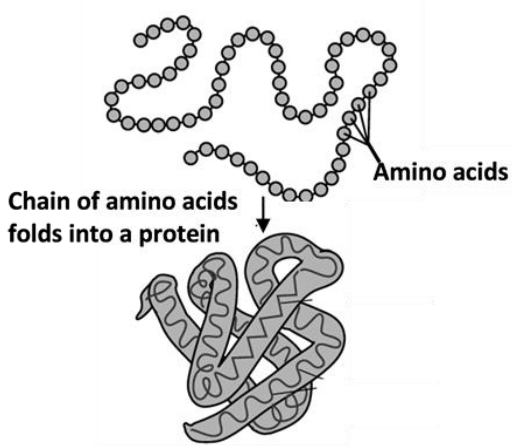


Each base in one strand of the DNA double helix pairs with a base in the other strand of the double helix. The base-pairing rules describe which bases pair together in a DNA double helix. Complete the following sentences to give the base-pairing rules.

3. A in one strand always pairs with _____ in the other strand.
C in one strand always pairs with _____ in the other strand.

Since all the nucleotides in DNA are the same except for the base they contain, each nucleotide is given the same symbol as the base it contains (**A**, **C**, **G**, or **T**).

A polymer consists of many repeats of a smaller molecule (a monomer). For example, a protein is a polymer of amino acids.



4. DNA is a polymer of ______________________________ .

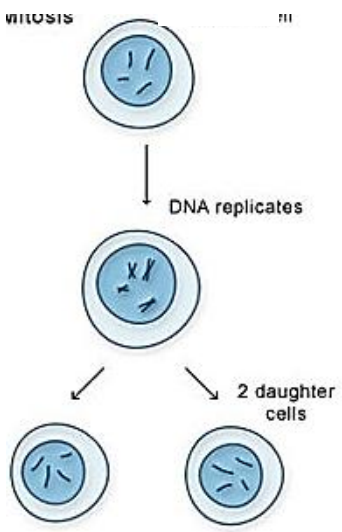
The sequence of nucleotides in a gene in the DNA determines which amino acids are joined together to form a protein. Slight differences in the sequence of nucleotides in a gene can result in different versions of the protein which in turn can result in different characteristics.

The sequence of nucleotides in a gene in the DNA
determines the sequence of amino acids in a protein which
determines the structure and function of the protein which
influences the characteristics or traits of the organism.

5. Explain how a difference in the sequence of nucleotides in a gene could result in one of these boys being albino and the other boy having normal skin and hair color.

DNA Replication

Our bodies need to make new cells to grow or to replace damaged cells. New cells are formed by cell division which occurs when a cell divides into two daughter cells. Before a cell can divide, the cell must make a copy of all its DNA; this is called DNA replication.

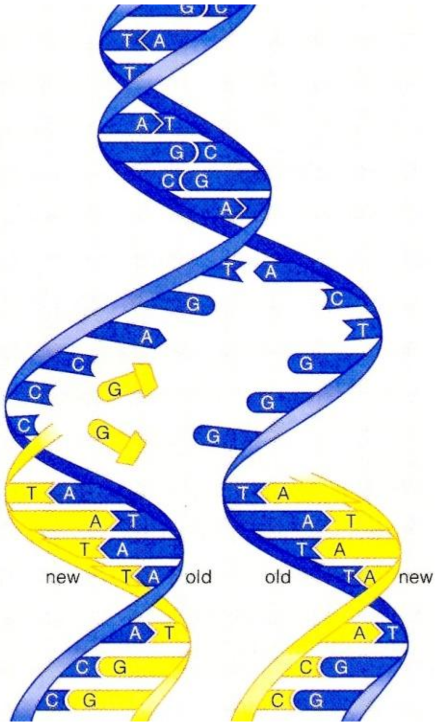


6. Explain why DNA replication is needed before a cell divides into two daughter cells.

During DNA replication, the two strands of the DNA helix are separated and each old strand provides the information needed to make a new matching strand. Each nucleotide in the new strand is matched to a nucleotide in the old strand using the base-pairing rules.

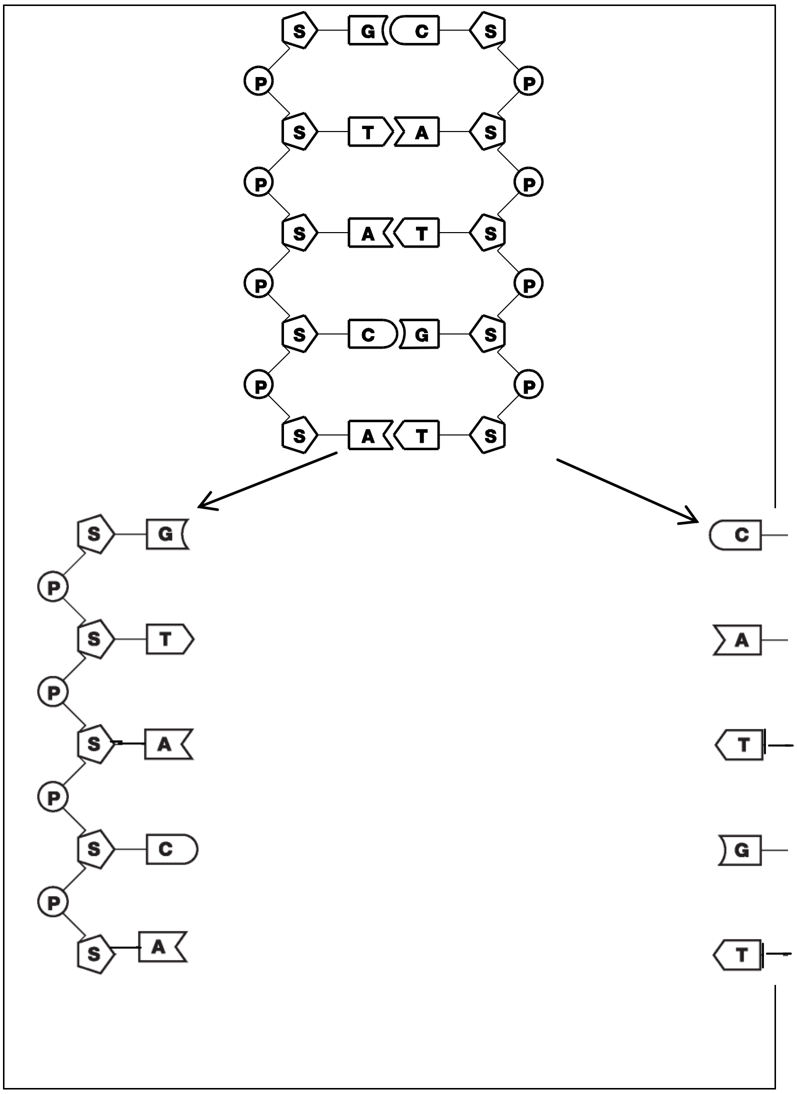
The enzyme DNA polymerase helps to make the new matching DNA strand by adding the matching nucleotides one at a time and joining each new nucleotide to the previous nucleotide in the growing DNA strand.

DNA replication results in two new DNA molecules that are identical to the original DNA molecule. Thus, each of the new DNA molecules carries the same genetic information as the original DNA molecule.



This drawing shows a short segment of DNA which separates into two strands in preparation for replication.

- Your job is to play the role of DNA polymerase and create the new matching strands of DNA to produce two pieces of double-stranded DNA. Add matching nucleotides one at a time, using the base-pairing rules and the nucleotides and tape provided by your teacher.



7a. Are there any differences between the two double-stranded pieces of DNA you have made?

7b. Are these new double-stranded pieces of DNA the same as or different from the original piece of DNA?

8. Why is it important that both copies of the DNA molecule have the exact same sequence of nucleotides as the original DNA molecule?

9. Based on the function of DNA polymerase, explain why each part of the name DNA polymerase (DNA, polymer, -ase) makes sense.

10. Explain how DNA polymerase, the double helix structure of DNA, and the base-pairing rules work together to produce two identical copies of the original DNA molecule.

From Gene to Protein – Transcription and Translation[i]

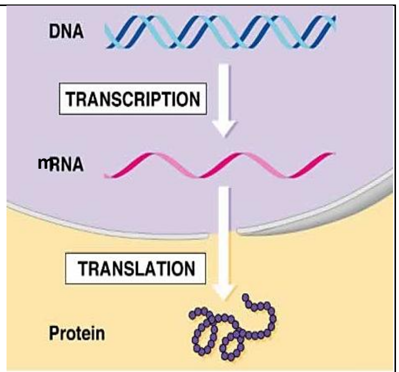
In this activity you will learn how the genes in our DNA influence our characteristics. For example, how can a gene cause albinism (very pale skin and hair)?

Basically, **a gene is a segment of DNA that provides the instructions for making a protein** and **proteins influence our characteristics**. This chart describes how two different versions of a gene can result in either normal skin and hair color or albinism.

DNA	→	**Protein**	→	**Characteristic**
	→		→	
Version of the gene that provides instructions to make normal protein enzyme	→	Normal enzyme that makes the pigment molecule in skin and hair	→	Normal skin and hair color
Version of the gene that provides instructions to make defective enzyme	→	Defective enzyme that does not make this pigment molecule	→	Albinism (very pale skin and hair)

The first step is **transcription** of the gene in the DNA. Transcription produces a messenger RNA (mRNA) molecule.

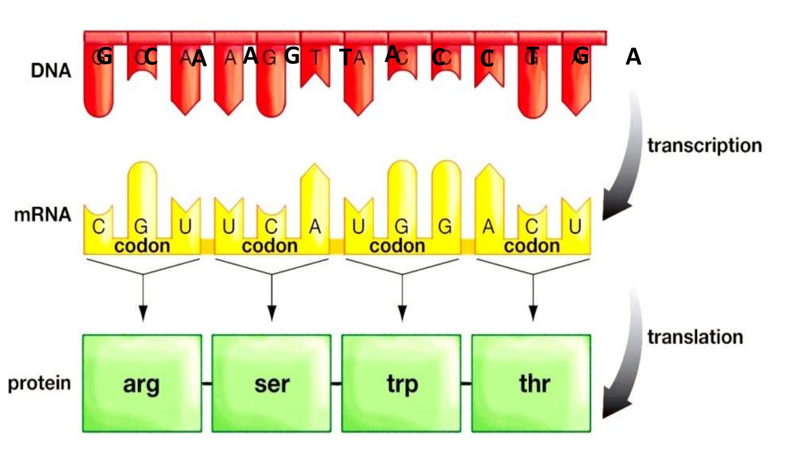
The second step is **translation** of the mRNA. Translation produces a protein molecule.



A gene directs the synthesis of a protein by a two-step process.

During **transcription**, the sequence of nucleotides in the gene in the DNA is copied into a corresponding sequence of nucleotides in mRNA.

During **translation**, the sequence of nucleotides in the mRNA determines the sequence of amino acids in the protein.



After translation, the sequence of amino acids in the protein determines the structure and function of the protein.

Notice that DNA and RNA are polymers of four types of nucleotides, **A**, **C**, **G**, and **T** for DNA and **A**, **C**, **G**, and **U** for RNA. In contrast, proteins are polymers of 20 types of amino acids.

1. To summarize how a gene directs the synthesis of a protein, label the process indicated by each arrow and fill in the blank with the appropriate molecule.

Gene in DNA ⟶ ____________________ ⟶ protein

2. Complete the following sentence to describe how differences in a gene can result in normal skin and hair color *vs.* albinism.

Differences in the sequence of ____________________________ in the gene

result in differences in the sequence of ______________________________ in mRNA which

result in differences in the sequence of _______________________________ in the protein which result in normal *vs.* defective enzyme to make the pigment in skin and hair which results in normal skin and hair color *vs.* __________________________.

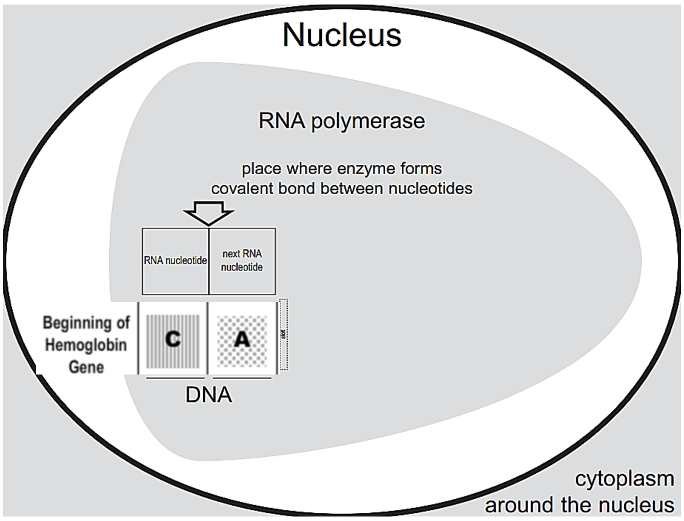
3. In this activity, you will model how a cell carries out transcription and translation to make the beginning of the hemoglobin molecule. What type of molecule is hemoglobin?

Transcription Modeling Procedure

Note: You will work with a partner to model the actual sequence of steps used by the cell to carry out transcription. You probably will be able to think of a faster way to make the mRNA, but you should follow the sequence of steps described below in order to learn how the cell actually makes mRNA. Remember, the goal is for you to simulate the actual molecular process of transcription in which the enzyme RNA polymerase carries out a step-by-step chemical process that adds one nucleotide at a time to the growing mRNA molecule.

- To model the process of transcription, you and your partner will need a page showing an RNA polymerase molecule inside a nucleus, a paper strip showing a single strand of DNA labeled "Beginning of Hemoglobin Gene", RNA nucleotides and tape.
- One of you will act as the RNA polymerase, and the other one will be the cytoplasm which surrounds the nucleus and supplies the nucleotides which are used to make the RNA molecule.

- RNA polymerase: Insert the "Beginning of Hemoglobin Gene" DNA molecule through the slot in the RNA polymerase diagram so the first two nucleotides of the DNA are on the dashes labeled DNA. Your RNA polymerase should look like this figure.

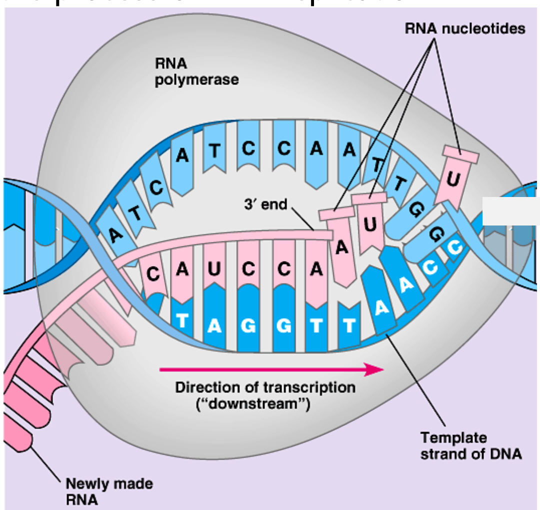


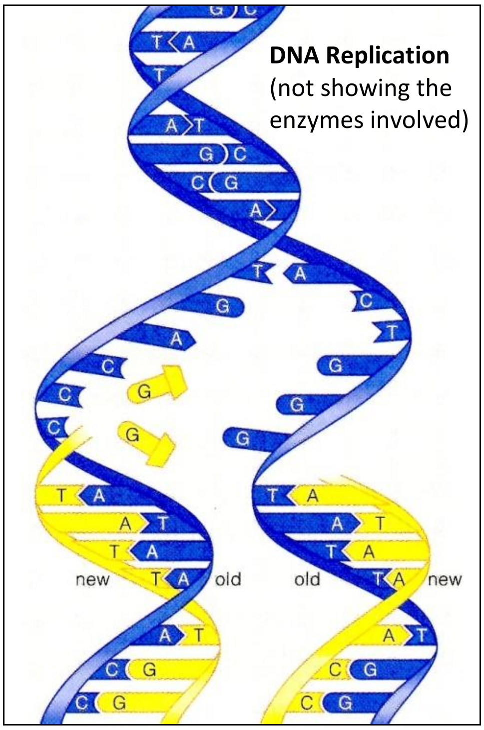
- Cytoplasm: Use the base-pairing rules to choose an RNA nucleotide that is complementary to the first DNA nucleotide. Give this nucleotide to the RNA polymerase person.
- RNA polymerase: Put the first RNA nucleotide in the box labeled RNA nucleotide.

- Cytoplasm: Give the next RNA nucleotide (complementary to the next DNA nucleotide) to the RNA polymerase person.
- RNA polymerase: Put this nucleotide in the box labeled "next RNA nucleotide" and join the two nucleotides together with transparent tape. The tape represents the covalent bond that forms between the adjacent RNA nucleotides as the mRNA molecule is synthesized. Then, move the DNA molecule and the growing mRNA molecule one space to the left.

➢ Repeat the last two steps as often as needed to complete transcription of the beginning of the hemoglobin gene, adding one nucleotide at a time to the mRNA molecule. Be careful to follow the base-pairing rule accurately, so your mRNA will provide accurate information for synthesizing the beginning of the hemoglobin protein when you model translation.

6. Describe three or more similarities between the process of transcription and the process of DNA replication.





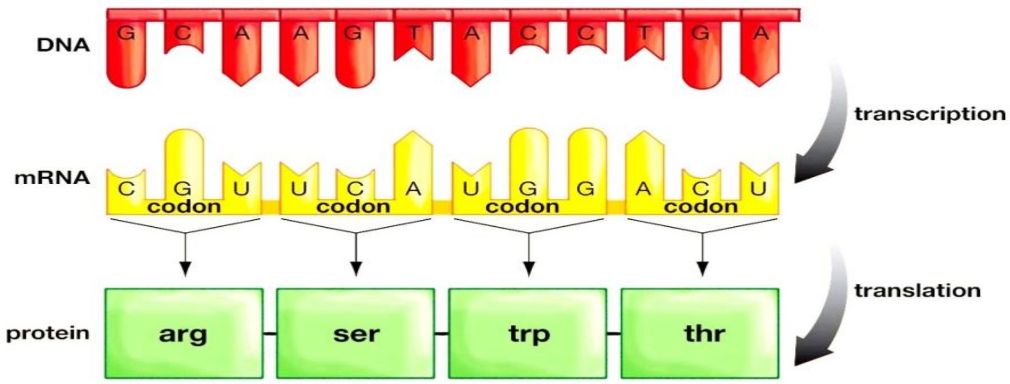
7. Fill in the blanks in this table to summarize the differences between DNA replication and transcription.

DNA replication	Transcription
The whole chromosome is replicated.	____________________is transcribed.
DNA is made. DNA is double-stranded.	mRNA is made. mRNA is ________________ -stranded.
DNA polymerase is the enzyme which carries out DNA replication.	_____ polymerase is the enzyme which carries out transcription.
T = thymine is used in DNA, so **A** pairs with **T** in DNA.	**T** = thymine is replaced by ___ = uracil in RNA, so **A** in DNA pairs with ___ in mRNA.

8. To summarize what you have learned about transcription, explain how a gene directs the synthesis of an mRNA molecule. Include in your explanation the words and phrases: base-pairing rules, complementary nucleotides, DNA, gene, mRNA, and RNA polymerase. Give your explanation in sentences and a labeled figure.

Translation

As you know, transcription is followed by translation. During translation, the sequence of nucleotides in mRNA determines the sequence of amino acids in a protein. Each set of three nucleotides in an mRNA molecule codes for one amino acid in a protein. This explains why each set of three nucleotides in the mRNA is called a **codon**. Each codon specifies a particular amino acid.

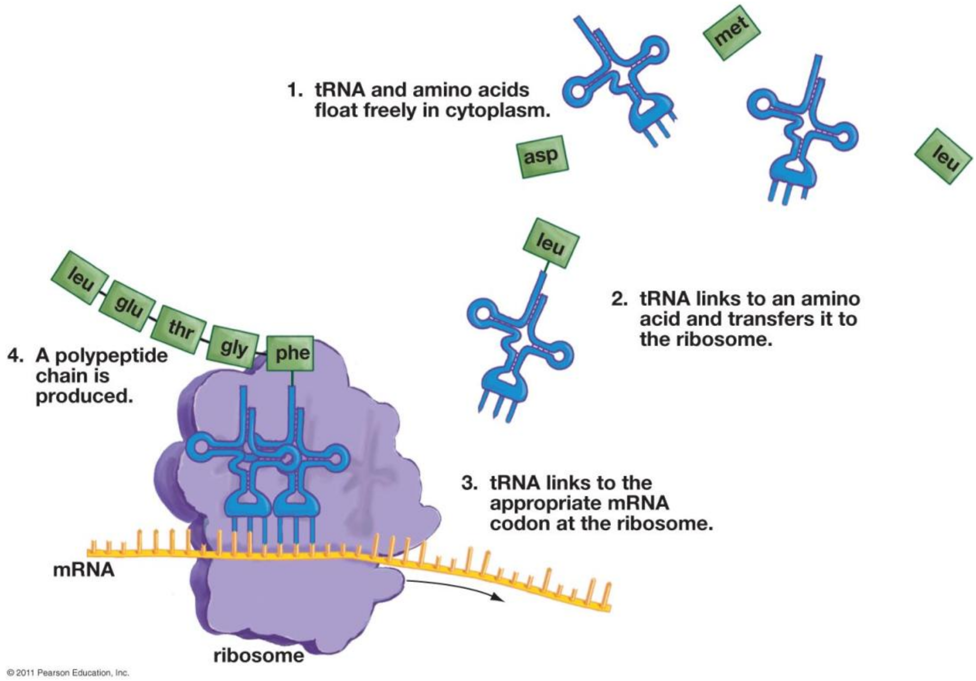


9. Fill in the blanks to complete the following sentence.

In the figure above, the first codon, _____, codes for the amino acid ____ (arginine) in the protein.

But how is translation accomplished in a cell? Translation is more complicated than transcription; the shape and chemical structure of each amino acid do *not* match the shape and chemical structure of the corresponding mRNA codon. Instead, a special type of RNA, **transfer RNA (tRNA)**, is required to ensure that the correct amino acid is brought in for each codon in the mRNA.

There are multiple different types of tRNA. Each type of tRNA molecule has three nucleotides that form an anti-codon. The three nucleotides in the **tRNA anti-codon** are complementary to the three nucleotides in the mRNA codon for a specific amino acid. For each type of tRNA, there is a specific enzyme that recognizes the anti-codon and attaches the correct amino acid to the tRNA (step 2).



ribosome is a tiny organelle where protein molecules are synthesized. The tRNA with amino acid enters the ribosome where the anti-codon in the tRNA is matched with a codon in the mRNA molecule (step 3). The tRNA brings the correct amino acid for that position in the growing protein molecule. Each amino acid is joined to the previous amino acid by a covalent bond (step 4). The ribosome moves along the mRNA, matching each codon with a complementary tRNA anti-codon and adding the appropriate amino acids one at a time to produce the protein coded for by the mRNA.

10. *Circle* the anti-codon in one tRNA molecule in the figure. In the ribosome, put a *rectangle* around an anti-codon in a tRNA and the complementary codon in the mRNA.

Translation Modeling Procedure

In this section you will simulate the steps in translation to produce the beginning of a hemoglobin protein.

- One of you will play the role of the ribosome and the other one will act as the cytoplasm, which is the source of tRNA and amino acid molecules.

Preparation:

- To prepare, you will need to have tRNA molecules, amino acids, the mRNA you made during your simulation of transcription, a strip labeled "Second Part of mRNA", and a page showing a ribosome. Tape the CUG end of the mRNA you made to the ACU end of the Second Part of mRNA strip.

- Cytoplasm: For tRNA molecules to function in translation, each tRNA must first be attached to the correct amino acid that corresponds to the anti-codon in that type of tRNA. To know which amino acid corresponds to each tRNA anti-codon, use the base-pairing rules to complete this table.

Amino acid	Threonine (Thr)	Histidine (His)	Proline (Pro)	Leucine (Leu)	Glutamic acid (Glu)	Valine (Val)
mRNA codon	ACU	CAU	CCU	CUG	GAG	GUG
Anti-codon in tRNA molecule that carries this amino acid	UGA					

- Cytoplasm: Use this table to match each model tRNA molecule with the correct amino acid for that type of tRNA. Tape the amino acid to the tRNA *very lightly*, because they will only be joined temporarily and will soon separate.

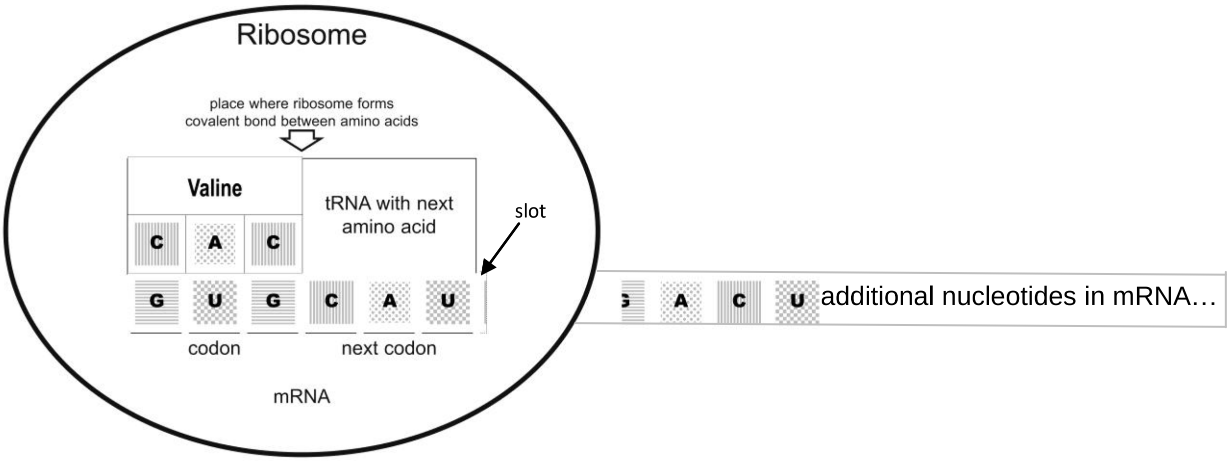
  Note: Each model tRNA molecule only shows the three nucleotides of the anti-codon and the binding site for the amino acid. A real tRNA molecule has many more nucleotides. Similarly, the mRNA molecule has many more nucleotides than shown in your strip.

11. Your partner wants to move ahead quickly, so he lays out the mRNA strip and puts the appropriate tRNA molecules above each of the six mRNA codons; then he tapes together all six amino acids. Explain why this would *not* be a good simulation of the actual sequence of steps used to carry out translation.

Modeling the Steps in Translation:

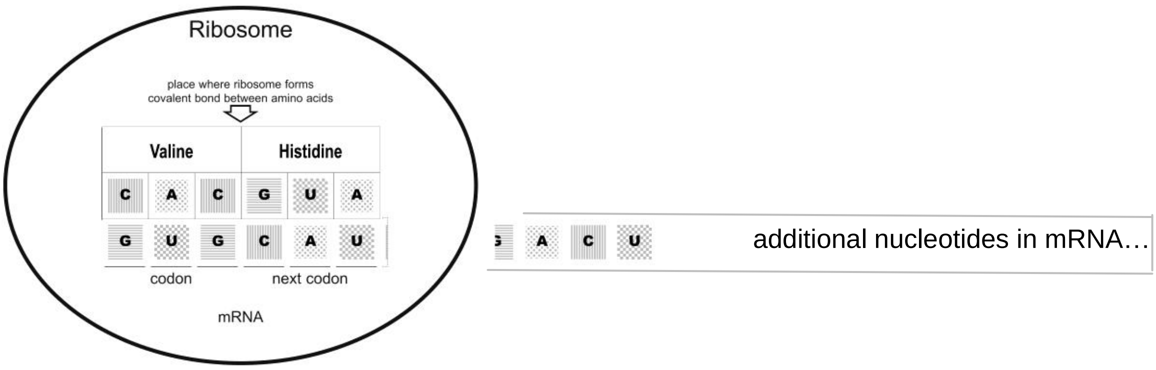
- Ribosome: Insert the mRNA through the slot in the model ribosome, with the first three nucleotides of the mRNA in the "codon" position and the next three nucleotides in the "next codon" position.
- Cytoplasm: Use the base-pairing rules to supply the tRNA that has the correct anti-codon to match the first codon in the mRNA.
- Ribosome: Place this tRNA with its amino acid in position.

Your model ribosome should look like:



12. In the above diagram, put a *rectangle* around each codon in the mRNA in the ribosome. In the tRNA, use an *arrow* to indicate the anti-codon, and use an * to indicate the amino acid.

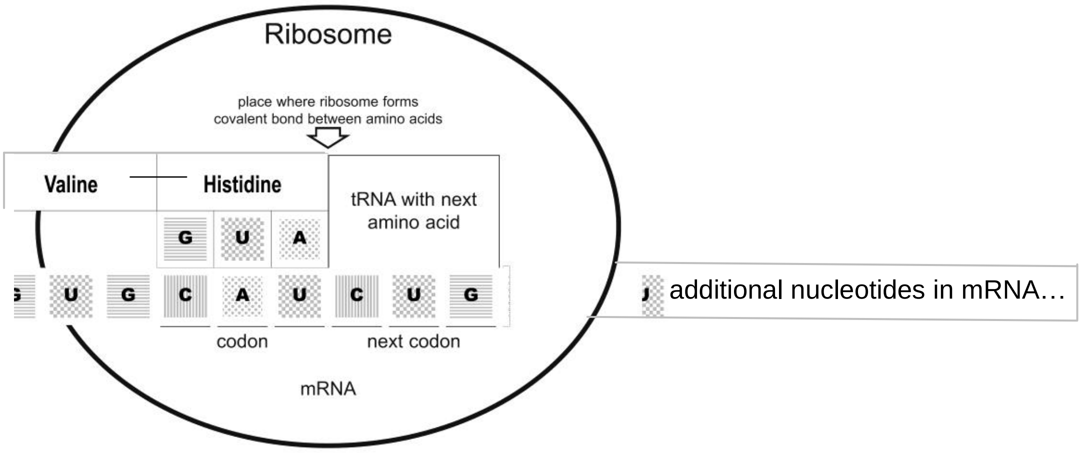
- Cytoplasm: Use the base-pairing rules to supply the tRNA that has the correct anti-codon to match the second codon in the mRNA.
- Ribosome: Place the tRNA in position. (Your model should look like the diagram below.) Now the ribosome is ready to link the first two amino acids with a covalent bond to begin the formation of the hemoglobin protein. Tape these two amino acids together; the tape represents the covalent bond between the first two amino acids in the hemoglobin protein. At this time, the first amino acid detaches from the first tRNA, so remove that tape.



13. *Draw a line* to indicate the location where you put the piece of tape to represent the covalent bond between the first two amino acids in the new hemoglobin protein that the ribosome is making.

- Ribosome: Move the mRNA to the left so the second codon is in the first position in the ribosome. At the same time, the matching tRNA with amino acid moves to the first position. Also, the first tRNA is released into the cytoplasm where it would be reused in a real cell.
- Cytoplasm: Put the first tRNA in the packet.

Your model should look like:



14. Why isn't the first tRNA shown in this diagram? What happened to it?

- Cytoplasm: Supply the tRNA that has the correct anti-codon to match the codon in the "next codon" position.
- Ribosome: Place the tRNA in position and tape the amino acid to the preceding amino acid. Then, move the mRNA and matching tRNAs with amino acids one codon to the left, and release the tRNA on the left to the cytoplasm person who will put it in the packet.

- Repeat this pair of steps until you have attached all six amino acids to form the beginning portion of the hemoglobin protein.

15. The proteins in biological organisms include 20 different kinds of amino acids. What is the minimum number of different types of tRNA molecules that must exist in the cell? Explain your reasoning.

16. What part of translation depends on the base-pairing rules?

17. Explain why a cell needs both mRNA and tRNA in order to synthesize a protein. Explain the function of mRNA, the function of tRNA, and how tRNA and mRNA work together to put the right amino acids in the right sequence as the protein is synthesized.

18. Explain why it makes sense to use the word translation to describe protein synthesis and why it would *not* make sense to use the word translation to describe mRNA synthesis.

19. Why does a cell need to carry out transcription before translation?

20. To summarize what you have learned about translation, explain how an mRNA molecule directs the synthesis of a protein. Include in your answer the words amino acid, anti-codon, base-pairing rules, codon, mRNA, protein, ribosome, tRNA, and translation. Give your explanation in sentences and a labeled figure.

____

Name ________________________ Date ____________ Period ________

Lab 12: Modeling Monohybrid Crosses

Background:

A **monohybrid cross** is a cross that involves one pair of contrasting traits. Different versions of a gene are called **alleles.** When two different alleles are present and one is expressed completely and the other is not, the expressed allele is **dominant** and the unexpressed allele is **recessive.**

Objectives: In this lab you will:

- Predict the genotypic and phenotypic ratios of offspring resulting from the random pairing of gametes.
- Calculate the genotypic ratio and phenotypic ratio among the offspring of a monohybrid cross.

Materials:

Lentils
Green peas
2 Petri dishes

Procedure:

Part I. Simulating a Monohybrid Cross

1. Write a definition for each of the boldface terms in the paragraph under Background:
   i. Monohybrid cross:

   ii. Alleles:

   iii. Dominant:

   iv. Recessive:

   v. Offspring:

2. You will model the random pairing of alleles by choosing lentils and peas from Petri dishes. These dried seeds will represent the alleles for seed color. A yellow lentil will represent *Y*, the dominant allele for yellow seeds, and a green seed will represent *y*, the recessive allele for green seeds.

3. Each Petri dish will represent a parent. Label one Petri dish "female gametes" and the other Petri dish "male gametes." Place one green pea and one lentil in the Petri dish labeled "female gametes" and place

one green pea and one lentil in the Petri dish labeled “male gametes.”

4. Each parent contributes one allele to each offspring. Model a cross between these two parents by choosing a random pairing of the dried seeds from the two Petri dishes. Do this by simultaneously picking one seed from each Petri dish *without looking.* Place the pair of seeds together on the lab table. The pair of seeds represents the genotype of one offspring.

5. Record the genotype of the first offspring in your lab report in Data Table A below.

6. Return the seeds to their original dishes and repeat step 3 nine more times. Record the genotype of each offspring in your data table.

7. Based on each offspring’s genotype, determine and record each offspring’s phenotype.

DATA TABLE A		
Gamete Pairings		
Trial	**Offspring genotype**	**Offspring phenotype**
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		

Part II. Calculating Genotypic and Phenotypic Ratios

8. Complete the data table below.

9. Determine the genotypic and phenotypic ratios among the offspring. First count and record the number of homozygous dominant, heterozygous, and homozygous recessive individuals you recorded in Table A. Then record the number of offspring that produce green seeds and the number that produce yellow seeds under "Phenotypes" in your Data Table B.

DATA TABLE B		
Offspring Ratios		
Genotypes	**Total**	**Genotypic Ratios**
Homozygous dominant (*YY*)		______: ______ : ________
Heterozygous (*Yy*)		
Homozygous recessive (*yy*)		
Phenotypes		**Phenotypic Ratios**
Green seeds		___________: ____________
Yellow seeds		

10. Calculate the genotypic ratio for **each** genotype using the following equation:

$$\text{Genotypic ratio} = \frac{\text{number of offspring with a given genotype}}{\text{total number of offspring}}$$

11. Calculate the phenotypic ratio for **each** phenotype using the following equation:

$$\text{Phenotypic ratio} = \frac{\text{number of offspring with a given phenotype}}{\text{total number of offspring}}$$

12. Now pool the data for the whole class, and record the data in your lab report in Data Table C.

DATA TABLE C		
Offspring Ratios		
Genotypes	**Total**	**Genotypic Ratios**
Homozygous dominant (*YY*)		______: ______ : ________
Heterozygous (*Yy*)		
Homozygous recessive (*yy*)		
Phenotypes		**Phenotypic Ratios**
Green seeds		__________: ___________
Yellow seeds		

13. Compare the class's sample with your small sample of 10. Calculate the genotypic and phenotypic ratios for the class data, and record them in your data table.

14. Construct a Punnett square below showing the parents and their offspring.

Part III. Cleanup and Disposal

10. Clean up your work area and all lab equipment.

11. Return lab equipment and materials to its proper place. Any materials taken from the front of the lab can be returns to its original bins in the front.

Final Analysis

1. **Summarizing Results** What character is being studied in this investigation?

2. **Drawing Conclusions** If a genotypic ratio of 1:2:1 is observed what must the genotypes of both parents be?

3. **Predicting Patterns** Show what the genotypes of the parents would be if 50 percent of the offspring were green and 50 percent of the offspring were yellow.

4. **Further Inquiry** Construct a Punnett square for the cross of a heterozygous black guinea pig and an unknown guinea pig whose offspring include a recessive white-furred individual. What are the possible genotypes of the unknown parent?

Mutation Activity:
What can happen when things go wrong?

Objectives:

i. To demonstrate the processes of transcription and translation.
ii. To demonstrate how the three types of mutations occur (insertion, deletion, and substitution).
iii. To demonstrate the effects of the three types of mutations on the amino acid chain produced by a DNA strand.

Background:
The genetic makeup of all known living things is carried in a genetic material known as DNA. The bases pair very specifically (A only with T and C only with G) so that when the DNA molecule replicates every cell has an exact copy of the DNA strand.

The order of the bases in a DNA molecule is the key to the genetic code of an individual. Every three bases are known as a codon and codes for an amino acid. Proteins are made up of amino acids and the order of them determines the protein made. In this way the order of the bases in the DNA molecule determines which proteins are made.

DNA is found in the nucleus of the cell, but proteins are made in the ribosomes in the cell cytoplasm. The mRNA molecule is used to carry the message from the DNA molecule in the nucleus to the ribosome in the cytoplasm. RNA is very similar to the DNA molecule except that the base T is replaced with the base U and RNA is single stranded (one half of the ladder).

At the ribosome, another type of RNA (tRNA) transfers amino acids from the cytoplasm to the growing amino acid chain at the ribosome.

BUT, sometimes there are problems with the DNA molecule that result in a change in the order of bases. This is known as a mutation and there are three different types.
1) Deletion: a mutation where a base is left out.
2) Insertion: a mutation where an extra base is added
3) Substitution: a mutation when an incorrect base replaces a correct base.

There are three possible outcomes when DNA sequences change:
1) An improvement
2) No change at all
3) A harmful change

Codon Chart

First Base	Second Base U	Second Base C	Second Base A	Second Base G	Third Base
U	Phenylalanine	Serine	Tyrosine	Cysteine	U
U	Phenylalanine	Serine	Tyrosine	Cysteine	C
U	Leucine	Serine	Stop	Stop	A
U	Leucine	Serine	Stop	Tryptophan	G
C	Leucine	Proline	Histidine	Arginine	U
C	Leucine	Proline	Histidine	Arginine	C
C	Leucine	Proline	Glutamine	Arginine	A
C	Leucine	Proline	Glutamine	Arginine	G
A	Isoleucine	Threonine	Asparagine	Serine	U
A	Isoleucine	Threonine	Asparagine	Serine	C
A	Isoleucine	Threonine	Lysine	Arginine	A
A	(start) Methionine	Threonine	Lysine	Arginine	G
G	Valine	Alanine	Aspartate	Glycine	U
G	Valine	Alanine	Aspartate	Glycine	C
G	Valine	Alanine	Glutamate	Glycine	A
G	Valine	Alanine	Glutamate	Glycine	G

Mutation Activity:
What can happen when things go wrong?

In this lab you will determine the protein for a normal strand of DNA and then the protein if each of the three types of mutations occurs for that particular strand of DNA.

Procedures:
1.The following is a strand of DNA that a protein will be made from. Write the "transcripted" mRNA in the spaces below it.
2.G – A – C – G – C – C – A – T – G – G – A – A – G – T – C
3.__ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __

4.Draw a line between each codon.
5. Look up the amino acid for each codon on the codon chart and write them in the spaces below. Be sure to do this in order. This is the "normal protein."
6.__________ - ___________ - __________ - ___________ - __________

7.The following is the same strand of DNA but with a deletion mutation in the second codon. Write the "transcripted" mRNA in the spaces below it.
8.G – A – C – G – C – A – T – G – G – A – A – G – T – C
9.__ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __
10.Draw a line between each codon. Do you see any differences between the codons on this mutated strand and the normal strand? __________
Describe them. ____
____
11.Look up the amino acid for each codon on the codon chart and write them in the spaces below.
12.__________ - ___________ - __________ - ___________ - __________
13. Was the number of amino acids the same as the original strand?
____
14. How many of the amino acids were the same as the original strand?
____
15. How many of the amino acids were different from the original strand?
______
16. Do you believe that this mutated DNA strand would create the same protein or a different protein as the original? ________
Why? ____
____

____

17.The following is the same strand of DNA but with an insertion mutation in the third codon. Write the "transcripted" mRNA bases below it.

18.G – A – C – G – C – C – A – T –A – G – G – A – A – G – T – C

19. __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __

20.Draw a vertical line between each codon. Do you see any differences between the codons on this mutated strand and the normal strand? _______
Describe them. ____
____

21. Look up the amino acid for each codon on the codon chart and write them in the spaces below.

22.__________ - ___________ - __________ - ___________ - __________

23. Was the number of amino acids the same as the original strand? _____

24.How many of the amino acids were the same as the original strand? _____

25.How many of the amino acids were different from the original strand? ___

26.Do you believe that this mutated DNA strand would create the same protein or a different protein as the original? ________
Why? ____
____

27.The following is the same piece of DNA but with a substitution mutation in the first codon. Write the "transcripted" mRNA bases below it.

28.G – A – A – G – C – C – A – T – G – G – A – A – G – T – C

29.__ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __ - __

30.Draw a vertical line between each codon. Do you see any differences between the codons on this mutated strand and the normal strand? _______
Describe them. ____
____

31.Look up the amino acid for each codon on the codon chart and write them in the spaces below.

32.__________ - ___________ - __________ - ___________ - __________

33. Was the number of amino acids the same as the original strand? ___________

34.How many of the amino acids were the same as the original strand? _____

35.How many of the amino acids were different from the original strand? ___

36.Do you believe that this mutated DNA strand would create the same protein or a different protein as the original? ________
Why? ____
____

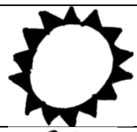
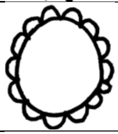
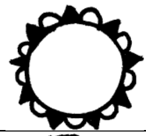
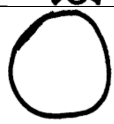
Using Blood Tests to Identify Babies and Criminals

I. Were the babies switched?

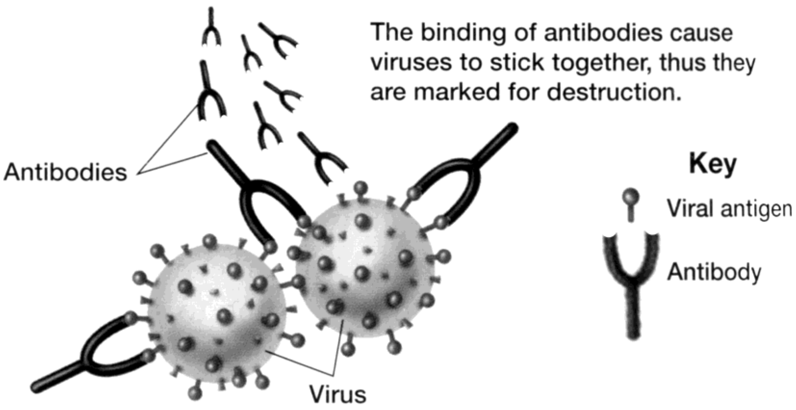
Two couples had babies on the same day in the same hospital. Denise and Earnest had a girl, Tonja. Danielle and Michael had twins, a boy, Michael, Jr., and a girl, Michelle. Danielle was convinced that there had been a mix-up and she had the wrong girl, since Michael Jr. and Tonja were both light-skinned, while Michelle had darker skin. Danielle insisted on blood type tests for both families to check whether there had been a mix-up. In order to interpret the results of the blood type tests, you will need to understand the basic biology of blood types.

Blood Types

There are many different ways to classify blood types, but the most common blood type classification system is the ABO (said "A-B-O") system. There are four blood types in the ABO system: Type A, Type B, Type AB, and Type O. These blood types refer to different versions of carbohydrate molecules (complex sugars) which are present on the surface of red blood cells.

People with:	**Have:**
Type A blood	Type A carbohydrate molecules on their red blood cells 
Type B blood	Type B carbohydrate molecules on their red blood cells 
Type AB blood	Type A and B carbohydrate molecules on their red blood cells 
Type O blood	Neither A nor B carbohydrate molecules on their red blood cells 

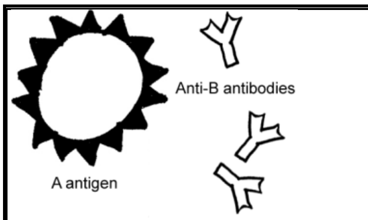
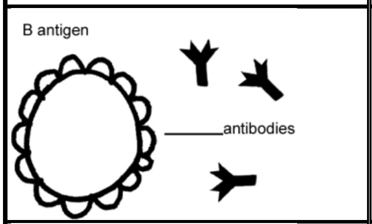
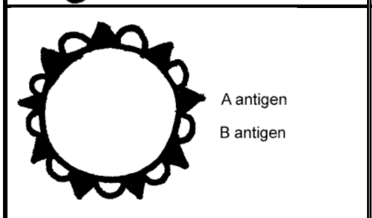
The Type A and Type B carbohydrate molecules are called antigens because they can stimulate the body to produce an immune response, including antibodies. Antibodies are special proteins that travel in the blood and help our bodies to destroy viruses or bacteria that may have infected our bodies.

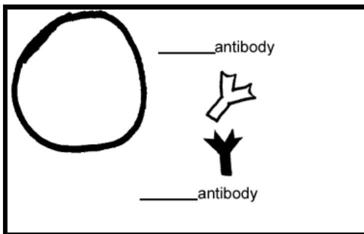


Normally, our bodies do not make antibodies against any molecules that are part of our own bodies. Thus, antibodies help to defend against invading viruses and bacteria, but normally antibodies do not attack our own body cells.

For example, people with Type A blood do not make antibodies against the Type A antigen which is present on their red blood cells. However, they do make antibodies against the Type B antigen (called anti-B antibodies).

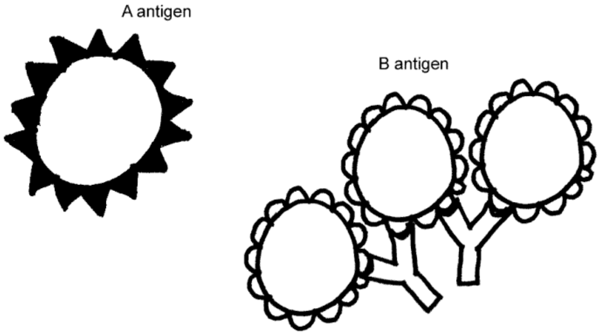
1. Test your understanding of blood groups by filling in the blanks in the chart below.

	**Blood group A** If you belong to the blood group A, you have A antigens on the surface of your red blood cells and ________ antibodies in your blood.
	**Blood group B** If you belong to the blood group B, you have B antigens on the surface of your red blood cells and ________ antibodies in your blood.
	**Blood group AB** If you belong to the blood group AB, you have both A and B antigens on the surface of your red blood cells and no anti-A or anti-B antibodies in your blood.

	**Blood group O** If you belong to the blood group O, you have neither A nor B antigens on the surface of your red blood cells, but you have both ________ and ________ antibodies in your blood.

Blood transfusions — Who can receive blood from whom?

If you are given a blood transfusion that does not match your blood type, antibodies present in your blood can react with the antigens present on the donated red blood cells. For example, if a person who has Type A blood is given a Type B blood transfusion, then this person's anti-B antibodies will react with the Type B antigens on the donated red blood cells and cause a harmful reaction. This transfusion reaction can cause the donated red blood cells to burst and/or clump together and block blood vessels.



Transfusion reactions can be fatal. To prevent this from happening, doctors test whether a person's blood is compatible with the donated blood before they give a transfusion. A person can only be given donated blood with red blood cells that do not have any antigen that can react with the antibodies in the person's blood.

2. Test your understanding of blood groups by completing the table below.

Blood Group	Antigens on red blood cells	Antibodies in plasma	Can receive blood from	Can give blood to
A	A	Anti-B	A and O	A and AB
B	B			
AB	A and B			
O	None			

3. Which blood type would be considered a universal donor (someone who can give blood to anyone)?

Genetics of Blood Types

The ABO blood types result from the alleles of a gene that can code for two different versions of a protein enzyme or an inactive protein as shown in this table:

Allele	Codes for a protein that is
I^A	a version of the enzyme that puts Type A carbohydrate molecules on the surface of red blood cells
I^B	a version of the enzyme that puts Type B carbohydrate molecules on the surface of red blood cells
i	inactive; doesn't put either type of carbohydrate molecule on red blood cells

4. Each person has two copies of this gene, one inherited from his/her mother and the other inherited from his/her father. Complete the following table to relate genotypes to blood types.

Genotype	This person's cells make	Blood Type
$I^A I^A$	the version of the enzyme that puts Type A carbohydrate molecules on the surface of red blood cells	
i i	the inactive protein	
I^A *i*	the version of the enzyme that puts Type A carbohydrate molecules on the surface of red blood cells and the inactive protein.	A

In a person with the IA i genotype, which allele is dominant, IA or i? Explain your reasoning.

5. Complete the following table to describe each of the three genotypes listed.

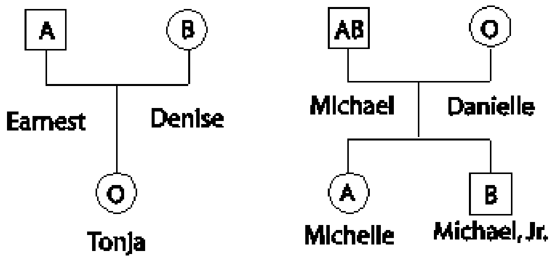
Genotype Will this person's cells make the version of the enzyme needed to attach this carbohydrate on the surface of his or her red blood cells?

Genotype	Will this person's cells make the version of the enzyme needed to attach this carbohydrate on the surface of his or her red blood cells?	Blood Type
$I^B I^B$	Type A __yes __ no; Type B __yes __ no	
I^B *i*	Type A __yes __ no; Type B __yes __ no	
$I^A I^B$	Type A __yes __ no; Type B __yes __ no	

Codominance refers to inheritance in which two alleles of a gene each have a different observable effect on the phenotype of a heterozygous individual. Thus, in codominance, neither allele is recessive — both alleles are dominant.

6. Which of the genotypes results in a blood type that provides clear evidence of codominance? Explain your reasoning.

Were the babies switched?
Now you are ready to evaluate whether Earnest and Denise's baby girl was switched with Michael and Danielle's baby girl. The following family trees show the blood types for each person in both families.



7. What allele for the blood type gene will be present in each egg produced by Danielle?

Michael can produce sperm with either the _____ allele or the _____ allele.

Draw the Punnett Square that shows the possible genotypes for Danielle and Michael's children. Write in the blood type for each genotype to show the possible blood types for Danielle and Michael's children.

Is it possible for Danielle and Michael to have a child who has type O blood?

8. To check whether Earnest and Denise could have a baby with Type O blood, draw a Punnett square for a father who has blood Type A and IA i genotype and a mother who has blood Type B and IB i genotype. Write in the blood type for each child's genotype.

9. Were Earnest and Denise the parents of Tonja or had the hospital made a mistake? Explain your reasoning.

10. How could fraternal twins be as different in appearance as Michelle and Michael, Jr., including one having light skin and the other having dark skin?

II. Who Killed Shamari Davis?

Background

Shamari Davis was a 20-year-old college freshman who was majoring in Physical Therapy. She paid for school by working as a personal trainer at a local gym. Shamari had been promoted to head personal trainer at the gym just before she was killed.

Crime Scene

The body was found in the women's locker room of the gym at 1 am by the night janitor, Harvey Willis. The victim had been strangled and was wearing a robe. There were signs of a struggle in the room and the glass door of the shower was broken and had traces of blood on it. The victim was pronounced dead at the scene and the coroner suggested that the time of death was at least 3 hours before the body was found.

Criminal Investigation

Shamari's co-worker Daleesha Jones told police that Shamari was a newer employee who did not deserve her recent promotion and only got it because she spent a lot of time with their boss, Steve O'Hare. When asked if he knew if Shamari had problems at work, Steve told Police that Shamari had complained to him that one of her fitness clients, Mike Reed, kept asking her out and wouldn't take no for an answer.

Blood Analysis

Obviously a real crime investigation would use many clues, but your investigation will be based on the simplest type of blood testing, namely testing for blood types A, B, O, and AB, for the blood sample found at the scene and for each of the possible suspects.

1. No individual can change their blood type, and blood type does not change with age. Explain why.

In order to test blood type, you mix a sample of the blood with two different types of antiserum—one which contains anti-A antibodies and one which contains anti-B antibodies. The reactions between the antibodies in the antiserum and the corresponding antigens on the red blood cells in the blood sample result in clumping.

2. Which types of blood have the antigens that will react with anti-A antibodies?

3. Which types of blood have the antigens that will react with anti-B antibodies?

4. Before you carry out the blood type tests, fill in the following chart that will help you to identify the blood type of each individual.

Reacts with anti-A antibody	Reacts with anti-B antibody	Blood type (A, B, AB, O)
Yes	Yes	
Yes	No	
No	Yes	
No	No	

Procedure

•Place your dish with the test wells on a piece of white paper, and put two drops of the blood of one suspect on both the A and B wells of the dish.
•Place two drops of anti-A antibody solution on the drop of blood in the A well and place two drops of anti-B antibody solution on the drop of blood in the B well.
•Mix the blood sample with the added anti-A antibody solution with one end of the toothpick. Mix the blood sample with the added anti-B antibody solution with the other end of the toothpick. Discard each toothpick after you use it.
•Record both reactions in the table on the next page, and record the blood type of the individual.
•Repeat this procedure for each blood sample.

	Reacts with anti-A antibody (Yes or No)	Reacts with anti-B antibody (Yes or No)	Blood type (A, B, AB, O)
Shamari Davis Victim			
Daleesha Jones Co-worker			
Harvey Willis Janitor			
Mike Reed Client			
Steve O'Hare Boss			
Blood on shower door			

5. Compare the blood types for the samples from the victim and each suspect to the blood type from the broken shower door glass at the scene of the crime. Use your observations to suggest who committed the murder.

Investigator's Report

6. Describe the circumstances which you believe led up to the crime, the time of the crime, and the individual that you believe is guilty of the murder. What evidence supports your conclusions?