Assessing Senior High School Science Students' Planning Skills In Selected Topics From The Integrated Science Syllabus

Anna M. Naah[*] **MPhil, Bilatam Peter Mayeem**[2]**M.Ed. Augustine Adjei**[3]**MPhil/M.Ed., Theophilus Aquinas Ossei-Anto PhD**

Abstract
The study was conducted to assess Senior High School Science Students' proficiency level in planning practical activities for 'Test of Practical' with selected topics from the Integrated Science syllabus. A total of 180 respondents (90 male and 90 females) out of a population of 1029 were sampled from three senior high schools in the Municipal Education, Offinso. Computer generated random numbers were used to randomly select 30 males and 30 females from each school in the municipality. The study employed the performed task instrument. Three tasks on planning were developed with their scoring formats and scoring details. Task A was on 'distillation'. Task B on 'density' and task C on 'osmosis'. A pilot test was conducted in one of the senior high technical schools in the Atwima Nwabiagya District Education and the response dichotomously scored, internal reliability of the various tasks was assessed using Kuder Richardson 20. The reliability of the various tasks obtained were task A, (0.731); B, (0.945); and C, (0.86). the data obtained for the actual study were statistically analyzed using descriptive, (percentages and means) inferential statistics (Independent t-test and Analysis of variance(ANOVA)). The students exhibited high proficiency on the planning skills. There were many similarities between the performances of female and male students of the various schools. The performance of students of school A were statistically significantly different from schools B and C on task B. from the major findings of this study, it is recommended that 'Test of Practical' questions should be related to real life situations

Background to the Study

Assessment in education may be described as the process of gathering, interpreting, recording, using and communicating information about students' achievements with respect to knowledge, concepts, skills and attitudes (Popham, 2000). Assessment is an important part of instructional programmes in schools (Williams, Howell and Hricko, 2006); as such assessment in science should be broaden to include performance based assessment that will provide a documentary record or a clear picture of students' abilities, interests and experiences or skills (Johnson, 2001).

Gaps in science achievement can be reduced by making science more relevant to the socio-cultural lives of students. Science, its methods of developing reliable knowledge and its applications in technology are at the heart of modern civilization (Seiler, 2001). However, the mode of teaching science in the Ghanaian context encourages memorization instead of understanding and application (Ameyaw-Akumfi, 2004; Buchele, 2008; Dowd, 2003; Fredua-Kwarteng & Ahia, 2005).

Basic scientific skills are needed for 'doing' science particularly at an early stage of the student's life. These basic skills include; the ability to raise questions and proposition of answers to such questions; ability to experiment or investigate; ability to find patterns in observations; ability to reason systematically and logically; ability to communicate findings; and the ability to apply what has been acquired in learning of science and other subjects (Keli, Haney, & Zoffel, 2009).

Performance-based assessment in science has been advocated as a means of assisting students in developing usable and transferable scientific skills. It also assists students in acquiring knowledge, understanding of the science epistemic (knowledge-generating) disciplines and easing the fear of entering into science programmes. The notion of performance practice was created after research had shown that

much of what people do in their everyday lives and on the job does not reflect the mathematics and science that they learned in school (Roth, Eijck, Hsu, Marshall, & Mazumber, 2009).

According to Shavelson, Baxter and Pine (1991), the call for alternative assessments for science achievement grows out of the current constructivist reform in science curriculum and cognitive research. Subsequently, performance assessments are generally valued for testing students' deep understanding of concepts and inquiry strategies, for making students' thinking visible, and for measuring skills in communicating about their science knowledge (Heidi, 2008). In addition, performance assessments can present authentic, real-world problems that can help students to apply academic knowledge to practical situations (Quellmalz, Hinojosa & Padilla, 1999).

One important hurdle science educators must overcome is getting students to have interest in science, so as to facilitate intellectual development of students, prepare and empower them to become more actively engaged in the decisions made in science. It is thought that students' attitudes toward science may have an effect on students' motivation, interest, and achievement in the sciences, gender notwithstanding. However, there are various biological differences in the make-up of humans as male and female which in most cases have led to the notion that one sex may have a 'learning edge' over the other sex (Mkpughe, 1998). Some people believe that because men are regarded as the dominant and even superior sex, they intrinsically have better brains and can learn much better than women (Mkpughe, as cited in Okoye, 2009). It has also been generally acclaimed that girls have better verbal ability than boys, whilst boys have greater visual spatial ability than girls. Perceived differences in ability of students (that is, learners) has often led to branding as noted by Ossei-Anto (1996), thus, "it is very easy to label students as non-performers in science because they tend to perform badly on paper-and-pencil laboratory assessment examinations" (p. 1) without taking due cognisance of the possible factors that work against them.

A research work on the use of Science Performance Assessment instrument was carried out by Ossei-Anto (1996) to assess laboratory skills of High Schools students in optics in Western New York. In the end, he asserts that Physics teachers need to expose their students more to non-traditional laboratory activities in order to develop and improve upon their skills of planning, performing and reasoning.

Several other researchers have conducted performance-based assessments and drawn interesting conclusions. Tachie (2001) used assessment-based performance tasks to assess observational skills of junior secondary school students and observed that students performed below average on observational tasks of classifying information when given specific rules. Anthony-Krueger (2001) developed performance-based assessment tasks to assess some process skills of senior secondary school students in elective Biology and concluded that "students could demonstrate some degree of interpreting, inferring and predicting skills" (p. 66). Addai (2001) carried out a study to evaluate practical skills of Senior Secondary School students using non-traditional tasks in mechanics and he concluded that non-traditional tasks are better able to measure students' laboratory skills. Similar study was conducted by Seshie (2001) to assess laboratory skills of students from selected senior secondary schools in Elective Chemistry on titrimetric analysis and observed that "the students could demonstrate some level of proficiency in the skills of planning and performing" (p. 48). However, despite the diversity of research on the teaching of science in Ghana using science performance assessment instruments to assess students' process skills, the researcher has not found any study reported on science performance assessment in relation to planning skills in Integrated Science at the SHS level. Therefore, giving rise to the need for this study.

Theoretical Framework
The science education profession works to expand its research base regarding student capacity to acquire scientific concepts effectively (North Central Regional Educational Laboratory, 2005). The science education literature states that shifting to an emphasis of active science learning requires a shift away from traditional teaching methods (National Academy of Science, 1996). According to (Myers & Dyer, 2004) the report by the American Association for the Advancement of Science emphasized that the teaching of scientific concepts should be consistent with the nature of scientific inquiry, which is fundamental to learning science. Myers & Dyer (2004) recommends that the process skill approach could be employed by science teachers in the effort to teach science as inquiry. This approach focuses on teaching broadly

transferable abilities that are appropriate to many science disciplines and are reflective of the behaviour of scientist (Padilla, 1990; Myers & Dyer, 2004).

The science process skills can be classified as either basic or integrated (Keli, Haney, & Zoffel, 2009). The basic science process skills serves as a foundation for learning the more complex integrated science process skills (Padilla, 1990; Myers & Dyer, 2004). According Myers & Dyer (2004), examples of integrated science process skills include skills such as formulating hypotheses, operationally defining, controlling, and manipulating variables, planning investigations, and interpreting data. Through this process, students engage in self-evaluation and set goals for their learning. They no longer serve as defenceless vessels waiting to be filled with facts. Instead, they are masters of their own learning and constructing their own meaning (Graves, 2002). This takes away focus from the teacher and lecture and puts it upon the student and their learning skills, including planning skills.Planning skill which is a Process skill in science can be achieved by students through performance practice. This could be done in carrying out the actual practical work, or giving a task for execution.

The study was premised theoretically on Bandura's Theory of Self-Efficacy (1997) and the Constructivist Theory (Teachnology, 2011). Self-efficacy is a person's beliefs concerning their capabilities to organize and implement actions necessary to learn or perform behaviours at designated levels (Janisch, Liu, & Akrofi, 2007). Although a person's beliefs about their capabilities are not the same as their actual ability, they are closely related; thus, if a person has a low self-efficacy regarding a certain task or concept, their performance in that area is expected to be low (Bandura, 1997; Myers & Dyer, 2004). Conversely speaking, higher ability levels would tend to increase self-efficacy levels and as a result increase the level of performance (Myers & Dyer, 2004). The constructivism learning theory argues that people produce knowledge and form meaning based upon their experiences (Teachnology, 2011) through assimilation which causes an individual to incorporate new experiences into the old experiences; this causes the individual to develop new outlooks, rethink what were once misunderstandings, and evaluate what is important, ultimately altering their perceptions; and accommodation, that is, reframing the world and new experiences into the mental capacity already present. Thus, individuals conceive a particular fashion in which the world operates and when things do not operate within that context, they must accommodate and reframe the expectations with the outcomes.

It is pertinent to note that the role of teachers is very important within the constructivism learning theory, where instead of giving a lecture the teachers in this theory function as facilitators whose role is to aid the student when it comes to their own understanding (Teachnology, 2011).A review of literature failed to identify research that has investigated the planning skills of students in the area of Integrated Science process skill at the level of the senior high school with the various school types and the influence of gender on planning skills. This information is needed in order to better assess the capability of students of senior high school offering Integrated Science, on planning skills. The findings from this study could be utilized by both Integrated Science teacher educators and by the students in science education in the development of planning skills regarding relating science concepts to solving real life problems.

Domains of Science

In science education, assessment is usually used to examine and describe student success and progress (Enger & Yager, 2001) in one or more of six domains of science: cognitive, psychomotor, application, affective, creativity and nature of science (Carin & Bass, 1997; Enger & Yager, 2001; Trowbridge, Bybee & Powell, 2000). The cognitive domain of science includes accepted scientific constructs such as scientific laws, principles and theories.

The psychomotor domain, often designated as performance or practical skills, includes science process skills such as: observing, manipulation of equipment/materials (assembling, measuring, and experimenting), planning, classifying, communicating, inferring, predicting, identifying and controlling variables, interpreting data, and formulating hypotheses (Dooley, Linder, & Dooley, 2005). The application domain requires the fortitude of the extent to which students can transfer what they have learned to a new situation, especially in their own daily lives.

The affective domain is mainly associated with explorations of human emotions, expression of personal feelings, decision making about personal values and about social and environmental issues (Dooley

et al., 2005). The creativity domain is essential to science as it is used by scientists in generating problems and hypotheses and in development of plans of action. Creativity calls for experience that promotes visualization (production of mental images), divergent thinking, consideration of alternative viewpoints, solving problems, and designing devices and machines. The domain on the nature of science is related to characteristics of science, knowing the world around us through empirical methods and how scientists think and work in the science community (Bell, 2009; Chabalengula, Mumba, Hunter & Wilson, 2009).

Assessment of skills in the psychomotor domain is directly associated with doing science in the laboratories. Generally, skills in the psychomotor domain can be manifested and demonstrated by students through hands-on activities in the laboratory (Rezba, Sprangue, Fiel & Funk, 1995). As such, this domain is important as it provides students with an opportunity to demonstrate their manipulation skills and understanding of processes and concepts through doing hands-on activities. The acknowledged weaknesses of conventional paper and pencil assessments have led to the recent development of alternative testing strategies (Slater, 1997). One of the most widely used of these is called performance assessment. The keystone of performance assessment is the use of a graded authentic task. An authentic task is one in which students are required to address problems grounded in real-life contexts. Such tasks are typically complex, somewhat ill-defined, engaging problems that require students to apply, synthesize, and evaluate various problem solving approaches (Shavelson et al., 1991; Oloruntegbe, 2010).

General Assessment

Assessment is a way of obtaining information about students and it includes the full range of procedures (observations, portfolios, ranges of projects, paper-and-pencil tests, oral presentation, exhibition and performance) that is used for making decisions about students, curricula activities and educational programmes (Parker & Gerber, 2002). According to Winking and Bond (1995), educators, policy makers and parents are beginning to recognize that minimums and basics are no longer sufficient for assessing students and are calling for a closer match between the skills that students learn in schools and the skills they will need upon leaving school. The situation whereby little or no skills are developed has also attracted a barrage of criticisms (Oloruntegbe, 2010; Ketelhut, 2007). A good assessment should be able to address this.

Assessment in education may be described as the process of gathering, interpreting, recording, using and communicating information about students' achievements with respect to knowledge, concepts, skills and attitudes (Kilfeather, O'Leary, & Varley, 2006). Assessment determines what, when and how students learn; additionally, assessment should enable a teacher have enough information about what students know, understand and can do and to help them learn (Onyango, 2008). In addition, the principle of assessment as an integral element of teaching and learning was espoused strongly in the Integrated Science syllabus (MOESS, 2007) as assessment could be used as a tool for enhancing students' achievement (Kilfeather et al., 2006).

The mode of assessment is changing for many reasons. Variations in the skills and knowledge needed for success; in appreciative of how students learn; and in relationship between assessment and instruction necessitates change in assessment modes. Thus, assessment modes should be tied to the design, content, new outcomes and purposes (Oloruntegbe, 2010). Of the science learning outcomes formulation of concepts, development of scientific culture and process skills and appropriate scientific attitudes, the cognitive area seems to attract greater attention and tend to be more handy and exciting in school teaching, learning and assessment. Thus, teachers seldom teach and assess skills and attitudes of students (Oloruntegbe, 2000; Oloruntegbe & Omoifo, 2000). The excessive use of paper-and-pencil multiple-choice tests has to be reviewed and the need for new and more varied assessment methods need to be emphasized (Downing & Haladyna, 2006). Assessments in which students carry out an activity or procedure to come out with a product in order to display their knowledge and skills are called performance-based assessments (Bekiroglu, 2008).

The use of performance-based assessment encourages teachers to employ inductive teaching by designing authentic tasks for learners to solve (Onyango, 2008). Such tasks resemble problems tackled by scientists; hence learners have to use scientific process skills to carry out their activities. The new method of assessment (performance-based assessment) focuses on the use of carefully constructed performance tasks

that give students the opportunity to exhibit and demonstrate, apply their skills and understanding as they would in the world outside the school

Performance - Based Assessments

The drive for change in science testing comes from the view that measurement of content knowledge, to the exclusion of process and application skills, gives an imperfect picture of students' science achievement (Jovanovic, Solano-Flores, & Shavelson, 1994). Besides, there is a belief that continual dependence on multiple-choice tests may hinder the development of innovative learning in students (Wiggins, as cited in Fairbairn, 2007). However, with performance–based assessment emphasis is placed on the process by which students generate solutions, and not just on the correctness or otherwise of the solution itself (Baxter, Shavelson, Goldman and Pine, 1992; Jovanovic et al., 1994; Kilfeather et al., 2006) and the belief is that individuals approach problem solving differently due to varying styles, and not differing abilities. This situation can be ameliorated by employing performance-based assessment. Thus performance-based assessment puts premium on problem-solving and critical thinking, and not guessing. This form of assessment also exposes students to science-related activities outside the classroom, which to a large extent gives opportunity to students to show what they know and can do (Buhagiar, 2007; Jovanovic et al., 1994).

Performance-based assessment can be described as an assessment which relies on the observation and judgment of activities as they occur' (Kilfeather et al., 2006). This type of assessment has three components: a task that requires students to solve a problem or to conduct an investigation using concrete materials in a hands-on way; a response format that allows students to communicate their findings; and a scoring system that allows judgment to be made about students' ability to carry out or complete a task. Performance-based assessment is sometimes regarded as synonymous to assessing real-life, with students assuming responsibility for self-evaluation. Testing is 'done' to a student, while performance assessment is done by the student as a form of self-reflection and self-assessment (Chabalengula et al., 2009). However, Brown and Shavelson (1996) point out that it is the addition of a scoring system that differentiates a performance assessment from a performance task.

Performance-based assessment is just one of the approaches that can be used in the classroom to gather information on students' progress and achievement. Few would dispute that it is the only type of assessment that should be used in the classroom or that it provides the most valuable kind of information about students. Performance-based assessment tasks are for assessing active learning experiences (Parker & Gerber, 2002). When students are engaged in activities that reflect the way professionals use and create knowledge in real-life contexts (Brann, Gray, Piety, & Silver-Pacuilla, 2010) the learning of science becomes meaningful to students. Thus, students should be engaged in learning activities similar to investigations of scientists and appropriate evaluation method should be used to assess students' knowledge and skills during investigative-type of activities. Research shows that well-designed performance assessments tasks put students' abilities, strengths and weakness in perspective ((Darling-Hammond, & Adamson, 2010).

Furthermore, performance-based assessments are tasks conducted by students that enable them to demonstrate what they know about a given topic. The difference between this type of 'test' and the traditional method is that students are given the chance to better communicate what they have learned (Lee-Ann, 2008). Students for example, may not be able to reproduce a lengthy definition; however, they can carry out an investigation and explain, in their own words, how and why it worked the way it did. Thus performance-based assessments afford students the opportunity to apply their knowledge by engaging in tasks that require critical-thinking strategies (Lee-Ann, 2008).

Assessment policies and practices in education are in a period of rapid transformation. The direct assessment of complex performances provides the impetus that is driving and guiding many of the current efforts to transform assessment (Gobert, Pallant, Krach & Daniels, 2010). Examples include a strong push to use more open-ended problems, essays, hands-on science problems, computer simulations of real world problems, and portfolios of student work. Collectively, such measures are normally referred to as 'authentic' assessments (Linn, Baker & Dunbar, 1991) because they involve the performance of tasks that are valued in their own right.

The movement of teaching science in the classroom to performance-based teaching is to provide equal opportunities for males and females to experience science in the classroom and outside the school. In

other words both males and females should actively perform in the performance-based science (Jovanovic & King, 1998).

The study of science is a critical element of science, technology, engineering, and mathematics (STEM) education. STEM learning scholars suggest that the most meaningful learning occurs when students are engaged in authentic activities that ask them to behave like professionals, for example, chemists, computer programmers, mathematicians, engineers, or archeologists. In this way, students engage in activities that reflect the way professionals use and create knowledge in real-life contexts (Herrington & Kervin, 2007; Tan, Yeo, & Lim, 2005). Advocates opine that performance assessment may be a more valued indicator of what students know and what they are able to do (knowledge and abilities) as it promotes active learning and deals with curricular-based testing (Shavelson, Baxter, & Pine, 1999).

Just like standardized achievement tests, performance-based assessments have norms, but the approach and philosophy are much different than traditional standardized tests. The underlying concept is that the student should produce evidence of accomplishment of curriculum goals. This can be maintained for later use as a collection of evidence to demonstrate achievement, and perhaps also the teacher's efforts to educate the student. The overriding philosophy of performance-based assessment is that teachers should have access to information that can provide ways to improve achievement. Also to demonstrate exactly what a student does or does not understand, relate learning experiences to instruction, and combine assessment with teaching (Kathy, 2010). The need to apply performance-based assessment as the focus for education reforms in assessment, curriculum and instruction was identified by researchers. Thus, to assess students on scientific reasoning and understanding rather than simply measuring discrete knowledge, critical assessment methods were developed, with a strong preference emerging for performance-based assessment (Morrison, McDuffie & Akerson, 2003; Scott, 2002). The leading proponents of this type of assessment (Akerson, Morrison and McDuffie, 2002; Guy & Wilcox, 2000; Perlman, 2003) argue that a performance-based assessment provides students with significant paths to exhibit their knowledge. The practice also helps to improve student skills by bringing into play complex functions of cognitive processing that involves a higher level of thinking for problem-solving, or the development of options when an individual confronts a new situation (Alsadaawi, 2008).

In addition, Spektor-Levy, Eylon and Scherz (2009) pointed out that performance-based assessment requires individuals to apply their knowledge and skills in context, not merely to complete a task on cue. Thus, performance-based assessments should be meaningful and engaging to students. With performance-based assessments, students are required to show what they can do, given an authentic task, which is then judged using a specific set of criteria.

Performance-based assessments go beyond measuring students' acquisition of knowledge. They demand far more than memorization of rules or facts. These authentic assessments aim to determine if students know how to apply their knowledge, demonstrating what they have learned through a variety of tasks (Darling-Hammond, & Adamson, 2010). Performance-based assessments are able to provide teachers with more detailed information than standard multiple-choice tests. They serve both a summative and formative purpose; they can tell teachers about what content a student has or has not mastered, and additionally offer insight into what concepts students are struggling with or where they get lost in a process (Darling-Hammond, & Adamson, 2010).

The characteristics of performance-based assessment make it imperative for students to engage with meaningful problems that foster significant educational experiences (Garbus, 2000). The performance-based assessment takes place over a period of time, and provides an opportunity for students to individually achieve the highest level of learning. Unlike the traditional testing procedures, performance-based assessment is a reliable assessment, because it involves the performance of tasks that are valued in their own right, it is situated in a real world context, and it can mirror actual tasks implemented by professionals (Brown, 2004).

The nature of performance assessment requires that the student demonstrate science process skills and knowledge in a practical setting. Typical science performance assessment furnishes students with laboratory equipment and poses a problem for students to solve (Klassen, 2006).

The development of performance assessments involves a general process. The processes involve three steps: defining the purpose, choosing the activity, and developing the scoring criteria (Wren, 2009).

The first step in developing performance assessments involves determining which concepts, knowledge, and/or skills should be assessed. The assessor needs to know what type of decisions will be made with the information gathered from the assessment. Secondly, the development of a performance assessment is to select the performance activity. The last step in constructing a performance assessment is developing the scoring criteria. Rubrics are used to evaluate the level of a student's achievement on various aspects of a performance task or product

The methodology of performance-based assessment in the classroom, provides teachers with timely information on the learning needs of their students (Corcoran, Dershimer & Tichenor, 2004). A research work conducted by Onyango (2008) to look at the introduction of performance based-assessment in a science classroom was to understand how alternative performance assessment can be embedded in science instruction to invite learners to engage in scientific inquiry. The process (the utilization of various assessment tools) was challenging and time-consuming to put into practice in the classroom. The researcher observed that, the enthusiasms that learners showed in class as they engaged in tasks was such a contrast to the previous passive sessions of listening and note-taking for future application on assignments and tests. As the students engage with the task through hands-on – minds-on activities, they learn to systematically solve problems presented to them; thus, acquiring scientific approach to issues and situations. Integrating performance assessment in classroom instruction makes students' performance the focus of the teaching and learning process and gives teachers the opportunity to teach according to what students already know and can do. This sort of situation is referred to as constructivist learning (Onyango, 2008). It also has value for students. For students, performance assessment provides a realistic approach to science as it reinforces the inquiry skills of science.

Performance assessment is therefore a suitable strategy for assessing students' concepts and skills in science, and it prepares students for a productive future within a technologically complex world (Alsadaawi, 2008). Importantly current goals for science educational standards reform present a significant shift to performance assessment (Atkin, Black, & Coffey, 2001). This is due to the fact that standards reform presents science as a subject where students are actively involved in science rather than reactive reading or listening (Alsadaawi, 2008).

An empirical study of the impact of performance-based assessment showed positive effects in the quality of students' learning and attitudes. Ainley, Hidi, and Berndorff (2002) observe that performance-based assessment does not only support the development of thinking and reasoning in the classroom, but also provides teachers with feedback that can be used to improve the classroom environment. Performance assessment is an appropriate strategy for assessing students' concepts and skills in science, and it prepares students for a productive future within a technologically complex world. A research conducted by Liu (2000) indicates that, students do not need to acquire a vast amount of information, typically the focus of traditional tests, but rather the ability to think, and organise that information for specific purposes.

A similar study conducted by Tüysüz, Karakuyu and Tatar (2010) to find out opinions of parents about performance task in science and technology class has revealed that parents are satisfied with the level of achievement about the performance tasks. Parents are of the view that performance tasks are useful and essential for the students, and contribute significantly to the students' social development. They prefer guiding their wards to do performance tasks on their own instead of doing it for them to find out the true results of their performance. They help their wards to control the phase of the tasks and to provide the equipment which are necessary, and also encourage them to complete their performance tasks. Thus, the study showed that there was a good collaboration between parents and teachers about students' performance tasks.

In another study, Biondi (2001) found out that performance-based assessment is a valid, equitable measurement of students' progress. Through performance assessment strategies, students become more focused on their work, and are able to reflect on their learning activities and develop a higher level of vocabulary through group conferences and self-assessments.

Many educationalists, however, propose that performance-based assessment should be considered not merely as a process for assessing students' understanding, but also as a learning process; one that teaches students concepts and requires them to explain and communicate their interpretations of the information, and their methodology for solving problems (Liu, 2000; Morrison et al., 2003). Hence, performance assessment methodologies and instructional objectives in science should be re-defined to include more practical

applications and more emphasis should be placed on synthesis and integration of content and skills. Therefore, a considerable change in instructional procedures and in science curricula, as well, must take place to align with theoretical conceptions that underline the new assessment method. In this situation, performance-based assessment can change classroom learning structures in which students merely listen and absorb information to those in which students actively participate by working together or separately. Furthermore, students in this learning experience, can assess their own progress and therefore, be more responsible for their own learning (Alsadaawi, 2008).

Performance-based assessment has a number of advantages be. It is important for gathering information on a wide range of learning expressions, underling concept acquisition, and development of psychomotor skills, developing communication skills and critical thinking. It also equips students with problem solving skills and esprit de corps and teamwork. It is one of the ways that teachers can use to assess the extent to which students can apply knowledge and skills to new situations. Moreover, it is useful for integrating assessment with teaching and learning to identify students learning needs, and for fostering pupil self-assessment (Kilfeather et al., 2006). Observing a student's work, rather than simply an aggregate score enhances the use of performance-based assessments, and it also offers teachers the opportunity to involve students more in their own learning and interests which include "reflection and expression of thinking processes (Tung & Stazesky, 2010).

A research that separated performance-based assessment from instructional procedures was conducted by Century (2002) who also compared the impact of alternative and traditional tests among sixth grade students. Century used the same teaching methods for both the control and the applied groups, but they were assessed differently using either performance-based assessment or a traditional test form and observed that there was no significant difference between students' performance on the two types of assessment.

Hammann, Phan, Ehmer and Grimm (2008) found out that performance test revealed a range of approaches to planning experiments among students, which are often flawed. A great number of students' strategies for planning experiments, are often unsystematic (Kuhn & Phelps, 1982; Hammann et al., 2008). Students have been shown to possess a non-scientific understanding of the aims and processes of experimentation. Consequently, students draw invalid conclusions, driven by confirmation bias (Chinn & Brewer, 1986), from their own ill-planned experiments (Hammann et al., 2008).

Science Process Skills

For students to develop scientific skills, it is important for them to be trained in the processes of seeking solutions to problems through scientific investigations and experimentations (MOESS, 2007). Scientific investigations and experimentations allow students to gain personal experiences of science through hands-on activities and to develop the skills associated with the practice of science (Cheng, 2008). These scientific investigations and experimentation are developed through the acquisition of science process skills. Process skills are used to describe a set of broadly transferable abilities that are reflective of what scientists do. They are fundamental to science, allowing everyone to conduct investigation, analyse data gathered, interpret them and draw conclusions. Process skills tend to last longer than learned content, and it is believed these thinking patterns can be readily transferred to new situations (Haigh, France & Forret, 2005). Science process skills include observing, inferring, predicting, controlling variables, hypothesizing, planning experiments and carrying them out. Studies in the United States have shown that elementary school students who are taught process skills, not only learn to use those processes, but also retain them for future use (Hofstein, Shore & Kipnis, 2004). In Ghana, the MOESS Integrated Science syllabus for Senior High Schools also emphasizes the teaching of basic process skills.

The syllabus design plays an important role in the acquisition of science process skills. The suggested time allocation guidelines of Integrated Science for Senior High Schools recommend an explicit teaching of the process skills; Year 1: Practical - Three continuous periods per week. Year 2: Practical - two periods as one double period. Year 3: Practical - 1 period per week. Year 4: Practical – one period per week, (MOESS; 2007, p vii).

Again, the profile dimensions for teaching, learning and testing in Integrated Science at SHS and their respective weights are as follows: Knowledge and Comprehension 20%, Application of Knowledge, 40% and Experimental and Process Skills 40%. Experimental skills involve the enquiry/investigative process of planning and designing experiments, carrying out case studies and field studies to be able to

compare phenomena or to observe phenomena closely to be able to identify causes, advance reasons for the occurrence of phenomena and develop practical solutions to problems and tasks. Process skills involve demonstration of practical manipulative skills using tools, machines and equipment for problem solving in science. Process skills also involve the processes of observation, classification, drawing, measurement, interpretation, recording, reporting, planning and expected scientific conduct in the laboratory/field (MOESS, 2007).

Science process skills are activities that scientists exhibit when they study or investigate a problem, an issue or a question. These skills are used to generate content and to form concepts. Rambuda and Fraser (2004) regard process skills as the way of thinking, measuring, solving problems, and using thoughts. This implies that thinking and reasoning are skills involved in investigative teaching and learning strategies and teachers and learners can apply science process skills while developing teaching and learning inquiry competences.

Virtually all educators agree that exposure to scientific investigations is an essential part of learning science (Kemi & Adsit, 2008). It is believed that online science courses, consisting of a thoughtfully designed sequence of investigations that are deeply interconnected with the relevant content instruction can provide this exposure equally as well (and sometimes better) than traditional classroom-based experiences (Kemi & Adsit, 2008). However, in both cases writing is crucial.

Many benefits of writing to learn in science have been established in research suggesting that writing supports students in developing the kinds of reasoning and communication skills. This scientific inquiry requires; as students write not only to represent their understanding to peers and teachers, but to construct their understanding of science content (Hand, Hohenshell & Prain, 2004; Hand & Prain, 2002). Karelina and Etkina (2007) conducted a study that focused on problem-solving in introductory college physics and found out that carefully–designed laboratory environments could improve students' ability to design an experiment to solve problems, collect and analyze data, evaluate assumptions and uncertainties, and to communicate. In a related study to ascertain, the role of scaffolding on students' laboratory experience, Karelina and Etkina (2007) observed that students whose use of scientific practices was scaffolded not only engaged in behaviours more like those of scientists than did students in traditional laboratories. However they also applied these behaviors in new situations where there was little or no scaffolding. Thus, as students learn content and scientific practices and have opportunities to apply those skills, they are better able to engage in novel tasks, more like what might occur in real-world problem–solving situation (Sutherland et al., 2010).

In another study, Etkina, Karelina, and Ruibal-Villasenor (2008) measured a variety of scientific abilities necessary to determine how assumptions might affect results, problem solving, such as students' ability to evaluate how experimental uncertainties may affect data. Also to identify the assumptions made in using a procedure, and to determine how those assumptions might affect results. They concluded that students' benefited from their engagement in the sequence of activities, the multiple cycles of investigative tasks, and their ongoing reflection.

The teaching and learning of science process skills are inseparable from the practice of science and play a key role in both formal and informal science content. Most jobs involve using these skills (Keli et al., 2009) and this makes science process skills not only important for those pursuing careers in science but for all. As a thorough knowledge of science content is impossible, mastery of science process skills enables students to understand, at much deeper level, the content they do know and equips them for acquiring content knowledge in the future (Keli et al., 2009).

Despite demands for performance assessment, many science curricula unfortunately, often over-emphasize content knowledge. But just as a literacy programme equips students with the basic tools of reading literacy, science literacy should also provide the tools required for all forms of scientific knowing (Colvill & Pattie, 2002). Furthermore, it is believed that content knowledge is acquired more efficiently and understood at a deeper level when obtained through inquiry using fundamental tools of science, the process skills (Keli et al., 2009).

Grumbine (2010) researched into the use of data-collection activities to enrich science courses by giving students the opportunity to collect, transform, and describe data as part of long-term scientific investigations that promote many positive outcomes. The study gave students a chance to experience the breadth and complexity of real-world data; they felt they were participating in real science, and they saw that their data added to the collective body of knowledge that future students would evaluate. This approach

allowed the students to practice and develop the necessary skills to uncover long-term trends or patterns that one-time-only data collection does not allow. It gave the students a sense of connection to the local environment around the school. Finally, the activities were engaging, fun, and motivating for students (and their instructors). Students need improved science process skills for their long–term academic and personal successes.

Assessing students' skills in experimentation is a challenging task. Planning an experiment requires successful employment of the control-of-variables strategy (Hammann et al., 2008). However, students have been shown to possess a wide range of approaches to planning experiments, and this diversity may not be adequately revealed by multiple-choice tests, which limit assessment to selecting between pre-planned experiments, rather than planning an appropriate experiment 'from scratch'.

Gender and Science Performance

Many research works have been conducted to assess process skills and efforts have been made to investigate students' performance along gender lines. According to Kohlhaas, Lin, & Chu, (2010) the term 'gender' refers to the social construction, usually based upon the biological make-up of one's body. Human gender can discriminate against males' and females' achievement in education. Shaw and Nagashima's (2009) research work on the achievement of student subgroups on science performance assessments in inquiry-based classrooms showed that girls outperformed boys on performance assessments. However, in their study, girls outperformed boys on both life (Ecosystems and Microworlds) and physical science (Food Chemistry) tasks. In a study using scores from four different performance assessments, Pine et al. (2006) observed comparable gender performance on physical science tasks while girls outperformed boys on the one life science task. Klein et al. (1997) observed that boys outperformed girls on a multiple-choice test. However, that same study found that girls outperformed boys on performance assessments.

On comparing gender in performance-based assessment, Ossei-Anto (1996) carried out a study on physics topics in optics at the State University of New York; Buffalo using students in selected high schools. He observed that males outperformed females on planning tasks whilst females outperformed males on performing and reasoning tasks. He concluded that the overall impact of gender on performance in science needs to be explored further. Johnson (2001) conducted a similar study using senior secondary school students offering biology at the Cape Coast and observed that females outperformed males on planning, performing and reasoning skills. Thus, it is clear that a generalization between males' and females' output in performance-based assessment could be premature as a large body of evidence is required before any such generalization can be made.

Mayer-Smitha, Pedrettib and Woodrowa (2000) researched on closing of the gender gap, using high school students enrolled in chemistry in North Carolina. They concluded that female students learning science, through and with technology, perform as well as or better than their male counterparts. Akpan (2002) also observed that, virtual laboratory environment provides equal chances for both genders. He concluded that, the finding should be investigated in detail with larger sample size. In a study conducted by Baser and Durmus (2010), gender was observed to have made a significant contribution to the variance in achievement related to direct current electricity concepts whereby males outperformed females. The achievement of males compared to females in respect of real laboratory environment may be attributed to the fact that males are generally more familiar with batteries, bulbs and wires than females. Ates (2005) had earlier found that achievement scores of male students were higher than those of female students during inquiry learning with batteries and bulbs when working on d.c. circuit.

However a research conducted on gender and racial/ethnic differences on performance assessments in science, females tended to score slightly higher than males on the performance assessments (Klein et al., 1997). It was noted however, that though certain types of performance task questions favored females, other types favored males. Klein et al. (1997) attributed this disparity to the emphasis a question places on certain cognitive abilities or skill experience. The findings suggest that differences in mean scores between males and females on performance measures were sensitive to the specific types of questions asked.

An important examination of the evidence from other studies suggests that there are more similarities than differences between the performances of females and males on practical laboratory tasks, despite well-established, sex-related differences in areas of interest, science-relevant experience and confidence (Ssempala, 2008).

In Nigeria, the seemingly poor performance of females in Integrated Science has been recognized (Ukwungwu, 2002) and great efforts have been made to study the influence of gender on performance in science. Unfortunately, these research efforts have not produced any definite clear-cut pictures from their findings as they do not agree on the magnitude and direction of gender differences in performance in Integrated Science (Ukwungwu, 2002). There is an abundance of relevant research cite, but often providing conflicting results and sparking academic controversies. Gender continues to inspire much on academic discussions. The present study tries to take a holistic view of gender effects on integrated science by looking at the effects of gender in relations to other variables of proficiency level, and school type

School Type

A school can be situated in an urban or a rural setting. Location of a school (rural or urban) affects a student's ability to study and perform at the level expected of him or her (Okoye, 2009). A stimulating school environment whips up students' interest to learn especially in the area of science. Hence the degree of interest that students derive from a learning environment affects their performance. Beaumont-Walters and Soyibo (2001), Mkpugbe (1998) and Okoye (2009) based school type on situation (rural or urban). Mkpughe (1998) noted that different aspects of school environment influence students' achievement. She further stated that the individual student's academic behaviour is influenced not only by the motivating forces of his home, scholastic ability, and academic values but also by the social pressure applied by the participants in the school setting. Thus, in carrying out a performance-based assessment, it is imperative that the school type is defined. School type has been variously defined by various researchers.

Ossei-Anto (1996) based school type on courses; Johnson (2001) based his on gender while Seshie (2001) based it on ownership (private and public schools). Ossei-Anto (1996) observed that the students offering Reagents Physics scored higher than those of General Physics. In Johnson's (2001) study, using mixed and single sex schools, girl schools outperformed the boy schools, whilst the boy schools outperformed the mixed school.

Combining school type and other variables on science performance may or may not produce the desired result (Okoye, 2009). Soyibo and Johnson (1998) also carried out a study on an analysis of high school students' performance on five integrated science process skills on school type and location of school. They observed that students could perform better when they receive better facilities and services of teachers of better quality. In this study, school type is defined as the background for the establishment of the schools in relation to a secondary/technical and SHS. The study therefore sought to assess students planning skills on selected tasks. Their performances were compared based on the school types (SHS offering General Science – School A; secondary/technical SHS – School B; and SHS. which is neither secondary/technical nor offer General Science – School C).

Research Design

Alternative assessment may be defined as any assessment format that is non-traditional, usually requiring students to exhibit activities such as construction, demonstration or performance. Alternative assessment formats or designs are more student-centred and authentic (Doran, Chan, Tamir & Lenhardts, 2002). Authentic is an assessment term that relates to "real world" situation or contexts, which generally requires a multiplicity of approaches to problem-solving and which takes cognizance of the fact that a problem might possibly have more than one solution. An example of alternative assessment is the performance-based assessment. The performance-based formats or designs are investigations, extended investigations and basic skills tasks (Doran et al., 2002).

Investigations

Investigations are at the core of an inquiry-oriented science course, especially one that employs the laboratory as a centre for science activities (Dora et al., 2002). Laboratory activities provide direct exposure of students to experience that reinforces knowledge and allows them to appreciate the investigative nature of science. In performance-based activity, students are required to analyse a problem, plan an experiment and execute it, collect data, organize and analyse them and communicate their findings. This approach enables students to experience and demonstrate their science inquiry skills and competencies by completing laboratory investigation. According to Doran et al. (2002), in some investigations, teachers can provide clues

to students if they are faced with difficulty at a particular step. This approach is acceptable as it is akin to the way scientists seek additional information from reference materials or colleagues when they are handicapped. However, students can be encouraged to seek reference materials from appropriate sources including internet websites in order to find their own clues.

Another approach to offering guidance is to organize investigations into a two-part format, with students carrying out and handing in the first part for review before continuing with the second part. For part 1, students just plan their investigations and submit their plan for review by teachers and peers. Students proceed with part 2 of the investigation by following their revised plan and carrying out an experiment they design themselves. On the other hand if their part 1 plan was not feasible, the teacher can provide a more workable plan. This ensures that all students are provided with an opportunity for success to be achieved. While this approach gives students less flexibility; it can present a safe, workable procedure enabling students to demonstrate what they are able to do.

Extended investigation
Extended investigations generally take place within a unit or lesson of a science curriculum, and are often associated with students' work on specific problems or projects (Doran et al., 2002). These forms of assessment are rooted in instruction, establishing a perfect fit between assessment and instruction. Extended investigation assessment format is the most natural and unobtrusive of the teaching-learning interface, because it forms part of teaching and learning experience in the science classroom. This format is the closet to instruction and its similarity to how problems are commonly encountered and addressed in real life. Students' work on extended investigations can be included in their portfolios (Doran et al., 2002; Chabalengula et al., 2009). It can be used to assess how well students are learning over an extensive period of time, rather than only their performance on an examination at the end of the lesson or unit. To develop hypotheses, plan experiment, solve problem and persist in reaching solutions can be achieved by students' ability to use an extended investigation. This can extended for days, weeks or even months.

Students can work independently and or in collaboration with others on an extended investigation. Assessment results of extended investigation can show students' persistence in ways that traditional testing methods cannot. With extended investigations time is allowed for students to show evidence of their planning and organizational skills, demonstrate their problem-solving skills as they carry out activities and demonstrate their skills at recording information and keeping records. An advantage of the extended investigation assessment format is that students can research into great depth a particular area of interest and can apply skills and knowledge learned in the classroom to a similar situation.

Basic skills tasks
Basic skills tasks centre on a narrow domain of skills. Basic skills tasks are short assessment (30 minutes or less) and usually focus on a small set of skills related to a particular situation or problem (Doran et al., 2002). Mostly science teachers refer to these tasks or assessment as, "station tasks", where students move from station to station; "bell ringer tasks", where a bell or other similar signal co-ordinates the movement of students from one task to another; "circus tasks", where student move in a circuit or circle; "partial inquires", where students complete one component of an investigation or laboratory experiment. Basic skills tasks focus on the assessment of skills in the psychomotor domain because it is directly associated with doing science. Generally, skills in the psychomotor domain can be manifested and demonstrated by students through hands-on activities (Rezba, Sprangue, Fiel & Funk, as cited in Chabalengula et al., 2009).

Basically basic skills tasks often require students to exhibit and demonstrate proficiency in manipulative skills such as measuring, using apparatus, and instruments, reading information from graphs, charts and tables, graphing and observing and following scientific procedure. Because basic skills tasks employed in-a-station, bell-ringer, circus and partial inquiry format and focus on a set of narrow domain skills, these assessment formats easily become part of activities within a unit of study (Doran et al., 2002). This study adopted the "basic skills tasks" format because a narrow domain of manipulative skills was assessed and also the period for the study was short. The other formats or design investigations and extended investigations were not use for the study because they require a whole unit of a lesson which could last as long as a whole term or year.

The station–by–station model was adopted and modified for use in the study. Tasks were explained on paper for students to go through. Tasks were placed on desks and students moved from one desk to other to plan task. Three tasks were involved which represented three stations. The various apparatus was set up in-front of the classroom. A total of 12 students were put on a task at a time for the pilot-tested study whilst a total of 20 students were put on a task at a time for the actual study. Students after completing each task moved to the next desk for the next task such that every student was able performs all the tasks.

Strength of the station–by–station format is that students are arranged in a manner in which students nearby each other will be performing different tasks, which makes the assessment reliable. Nonetheless, the stations model is not a panacea for assessing manipulation skills during science labs as the stations are an artificial construct that separates assessment of lab skills from the performance of the activity itself, (Harden & Cairncross as cited in Chabalengula et al., 2009).

Population

Population consists of all subjects (could be human or otherwise) that are being studied (Bluman, 2004). All final year senior high schools were the target population. These schools have been classified into four (A, B, C, and D) by the Ministry of Education. The accessible population for the study was senior high school students of category 'C' because it had the various school types under study and the subjects were offer Integrated Science as a core subject.

The population for the study comprised of final year students from three Senior High Schools in the Offinso Municipal District. The schools belong to Group C of the Ministry of Education classification of schools. Schools in this category have similar infrastructure, performance and offer Integrated Science as a core subject. The schools selected were: a SHS offering General Science (school A); a secondary/technical SHS (school B); and a SHS which is neither a secondary/technical nor offer General Science (school C). The population size drawn from the schools was 1,029. The population for school 'A' was 367 (202 males and 165 females), school 'B' was 253 (129 males and 124 females) and school 'C' 509 (264 males and 245 females).

Sample and Sampling Procedure

Sample is a group of subjects selected from a population (Bluman, 2004). There are two methods used in selecting the sample. These are probability sampling and non-probability sampling. With probability sampling all the subjects have an equal chance of being selected. The methods employed in probability sampling include: simple random sampling, where subjects are selected by random numbers; systematic random sampling, where subjects are selected by using every 'kth' number after the first subject is randomly selected from 1 through 'k'; stratified random sampling, where a population is divided into subgroups, called the strata and subjects are selected from each stratum and cluster sampling, where the subjects are selected by using an intact group that is representative of the population (Bluman, 2004). For non-probability method not all the subjects have an equal chance of being selected. The methods employed in non-probability sampling include purposive sampling and convenience sampling (Mason, Lind, & Marcahal, 1996). Purposive sampling subjects are selected based on their knowledge on the topic under study. With convenience sampling employs subjects that are convenient and available.

Various authorities have come up with a variety of formulas for determining sample size under different sampling methods (Bluman, 2004; Sarantakos, 1988). However, for this study, none of these formulas was used due to the following reasons; it was not possible to use a larger sample size due to non-availability of sufficient apparatus to go round all the students, it was not possible to use a larger sample size in each school for fear of contamination whereby students could have prior information of what their colleagues had done before they themselves were engaged in the study. The schools made only a time available for the study due to their tight schedules. The sample size for the three schools of the study was 180 students which comprised 90 males and 90 females. This was considered adequate in view of the fact that for a sample size of 30, the shape of the sampling distribution of the mean approximates the normal distribution (Hill & Lewicki, 2007).

In each school, final year students were assigned unique numbers respectively. Simple random sampling method was used to select 60 students from each of the school. This was to give all the students an equal chance of being included in the sample for the study. Computer–generated table of random numbers

was used to select 30 males and 30 females. The average age of the students was 19 years. The students have had 9 years of basic education made up of 6 years of primary school and three years of J.H.S education and are in SHS final year preparing for the West African Senior Secondary Certificate Examination (WASSCE).

Instruments

Instruments are tools researchers use to collect data for research studies (alternatively called "tests"), Jacobs (2010). These include questionnaire and interview. A questionnaire is a series of questions asked to individuals to obtain statistically useful information about a given topic. Questionnaires usually are comprised of a number of different approaches to asking questions – the essential ones being: closed questions, multiple choice and open-ended questions (Biringham & Wilkinson, 2003). Interviews are a systematic way of talking and listening to people and are another way to collect data from individuals through conversations (Kajornboon, 2005). Interviewing is a way to collect data as well as to gain knowledge from individuals. It is regarded as interchange of views between two or more people on a topic of mutual interest. There are many types of interviews, which include: structured interviews, semi-structured interviews, unstructured interviews, and non-directive interview (Kajornboon, 2005).

Besides these two, various methods have been used to assess students' achievement in science lab activities. Some of these methods include: written lab reports in a form of open-response question, or multiple-choice, performance tasks, portfolios, and self/peer checklist and investigative projects (Chabalengula, Mumba, Hunter & Wilson, 2009). Performance tasks involve students demonstrating their understanding through actual manipulation of equipment and materials in the laboratory. Performance task creates to some degree the conditions in which scientist work and to elicit the kind of thinking and reasoning used by scientist to solve problems (Yujing, 1997). The study used the developed performance task as instrument to assessed students planning skills

By way of instruments, three tasks were designed for the study based on concepts derived from the Integrated Science (test of practical) syllabus. The concepts were used to develop tasks because they lend themselves to application in everyday activities and offer explanation to some phenomena such as transport in organisms, measuring densities of irregular objects and water purification process. Developing tasks such as these that reflect on out-of-school situation motivate students, generate interest and help put science into action in a more meaningful, non-complex situation (Addai, 2001). The three tasks were in the domain of planning. Each of the tasks was unique, complete and independent and contained different basic skills that ensure that students' difficulty in any one the tasks was not carried over to the next task. Scoring formats were developed for each task.

Opinionnaire was used in the pilot-test. Students gave their opinion on various tasks they planned based on difficulty of tasks, familiarity with apparatus, the relatedness of concepts and unfamiliar words they will come across. Their opinions were then used to revise the instruments for the actual study (Opinionnaire was not used in the actual study).

Performance Task

The tasks were on planning. The tasks were developed by explaining the meaning and the principle of the concept concerning each activity. A problem relating to real-life situation was posed in relation to the concept employed. Diagrams were drawn to illustrate the tasks. Students were also provided with a list of materials and were required to indicate how they would use these materials. The actual materials for the planning the various tasks were provided on the day of administering the tasks (in case some students would want to confirm their procedure through performance of actual experiments). Students were provided with answer booklets where they outlined the steps/procedures (Appendix H). Subsequently, scoring formats and scoring details (for clarification of scoring) for the various tasks were developed and used in scoring the tasks. The scoring formats contained the various levels of planning skills that should be exhibited by the student on each task (Appendices E, F, G and H).

Pilot-test of Task, and Opinionnaire

The tasks (A, B and C) and opinonnaire were pilot-tested for the feasibility rates of reliability and validity of the instrument. Final year students of one of the secondary/technical SHS of the Atwima Nwabiagya District were used for the pilot-test. These students had similar characteristics and also offer Integrated Science course as the students that were involved in the actual study. Before the pilot-test experts in the Department of Science and Mathematics Education of the Faculty of Education, University of Cape

Coast, read through the tasks and opinionnaire to examine the face and content validity in order to ensure that they were devoid of ambiguities.

A pilot-test was conducted on 8th November, 2010 at one of the secondary/technical Senior High Schools (co-education) in the Atwima Nwabiagya District in the Ashanti Region. The school was selected because it belongs to Group C of the Ministry of Education classification of schools. Schools in this category have similar infrastructure, performance and offer Integrated Science as a core subject. This is particularly important since the intended schools for the actual study belong to the same category. Thirty-six students, made up of 18 males and 18 females, were randomly selected using computer-generated table of random numbers from a population of 405 students; comprised of 180 females and 225 males. Responses of students used in the pilot-test were scored using a scoring format (Appendices D, E and F). Scores were subjected to complete item analysis using reliability analysis of SPSS, Version 16 to determine the variance and Kuder Richardson (KR 20) to determine the internal consistency reliability as follows:

$$KR20 = \frac{k}{k-1}\left(1 - \frac{\Sigma pq}{\sigma^2}\right)$$

where k = the number of items in the assessment, instrument

σ^2 = variance,

p = proportion of students who had an item correct and

q = proportion of students who had an item wrong.

The Kuder Richardson (KR 20) formula was used because the tasks were dichotomously scored (0 or 1). The reliability of the various tasks obtained were as follows; task A (0.731), B (0.945) and C (0.860). Student impressions were solicited on tasks relatedness, difficulty of tasks, unfamiliar words they have come across and their familiarity with materials provided. The responses from students were used to revise the instrument which was then used for the actual study. Some words that students were not familiar with were replaced (example "enskinment" replaced with enstoolment, "phenomenon" with process).

Data Collection Procedure

Three schools (A, B, and C) were used in the study. The schools are in the Offinso Municipal District. The Offinso Municipal District was chosen because of familiarity and the rapport between the researcher and heads of the institutions and science teachers. The Offinso Municipal District of Education has three senior high school types (SHS offering General Science, Secondary/technical SHS and SHS that neither offer General Science and also not a Secondary/technical) under one category (Group C). The researcher made inter-personal contacts with the heads of the three SHS(s) that offered Integrated Science. Formal letters of request obtained from the Department of Science and Mathematics Education, Faculty of Education, University of Cape Coast were personally taken to the heads of the various school types. After approval by the heads of the schools, the heads of science department were contacted through the assistant headmaster (academic) of the various schools. Arrangements were made with the heads of science department to obtain students' population, their class lists (for sampling), and information on the dates and times that were suitable for data collection. From the information gathered schedules were worked out for the final administration of the tasks. Two schools (A and B) were visited on the same day as agreed upon. The researcher met the heads of science department and followed this up by meeting with the students who were sampled for the study. The students were informed about the reasons for carrying out the exercise. The students were assured of their individual and group confidentiality. Their maximum co-operation was solicited to guarantee the success of the study.

Data for the pilot-test was collected on 8th November, 2010. Thirty-six students were arranged in 6 x 6 columns and rows. Student 1 was given task A, student 2 task B and student 3 task C. The same sequence was followed till every student was assigned a task. The first student of the next column was given task B, the second student was given task C and task A was given to the third student. The order was followed till all of the students were covered. An apparatus was provided for the each task. Students took turns to perform the same task. At any given time every student was engaged in a task. The answer sheet of each student was collected after completing each of the three tasks.

The actual data was collected over 2 days, on 22nd and 23rd November 2010. Two of the schools (A and B) were visited on the 22nd November, 2010. Schools A and B were visited on the same day. The event for School A took place before breakfast while that of School B occurred after second break in accordance

with schedules that were provided by the heads of science departments. The third school, C was visited on the 23rd of November 2010. The heads of department helped with the arrangements of the class for the data collection. The data was collected by the researcher. Similar to what was done at the pilot-tested study, the students were arranged according to a 6 (columns) x 10 (rows) matrix. Student 1 was given task C, student 2 task B and student 3 task A. The same sequence was followed till every student was assigned a task. The first student on the next column was given task B, the second student was given task A and the task C was given to the third student. The order followed till all of the students were covered. An apparatus was provided for the each task. Students took turns to perform the same task. At any given time every student was engaged in a task. The answer sheet of each student was collected after completing each task before he or she moved to tackle the next task.

Data Analysis

The answer sheets containing responses of the students were scored using scoring format. This was done by two raters. Students answer sheets were coded using numbers. After coding, the scores were entered into the computer. Inter-rater reliability was computed using Pearson correlation. The inter-rater reliability of the various tasks was obtained as followed; task A (0.904), (0.914) and (0.903) for schools A, B and C respectively; task B (0.917), (0.909) and (0.918) for schools A, B and C respectively and task C (0.913), (0.901) and (0.911) for schools A, B and C respectively.

Both descriptive and inferential statistics were carried out on the data using the Statistical Package for Social Scientists (SPSS) software Version 16 and the Microsoft Excel 2007 software. Descriptive statistics was used to analyze the data collected for research question one ("To what extent do students exhibit planning skills?"). Percentages scores obtained by students were calculated using the SPSS version 16 software. Microsoft Excel 2007 software was used to draw graphs to determine the level of proficiency exhibited by students. Inferential statistics (independent t-test and ANOVA) were used to analyze research questions two ("Which gender shows more proficiency in planning skills?") and three ("Which school type exhibits most proficiency level in planning skills?"). Independent t-test was used to analyze the data collected for research question two, thus comparing the performance of male and female students on the various tasks in the various schools. ANOVA was used to compare the school type and means performances of the various tasks of the various school. A post-hoc comparisons test was done using Turkey's Honestly Significant Different (HSD) to identify where the various differences occur and Eta square was used to calculate the effect size to determine the relative magnitude of the difference between the means.

Results

The results were organized and presented under the research questions as follows:

**Research Question One: To what extent do students of the SHS offering Integrated Science exhibit planning skills?**

General Details
(a) Task A: (Appendix A and Appendix D)

Generally, proficiency was higher than 75% for all the skills exhibited for task 'A' of school A (Figure 1). On the whole, the students scored more than 90% in three skills of the proficiency levels.

Proficiency level
Figure 1. Assessment of planning for task 'A' of school 'A'
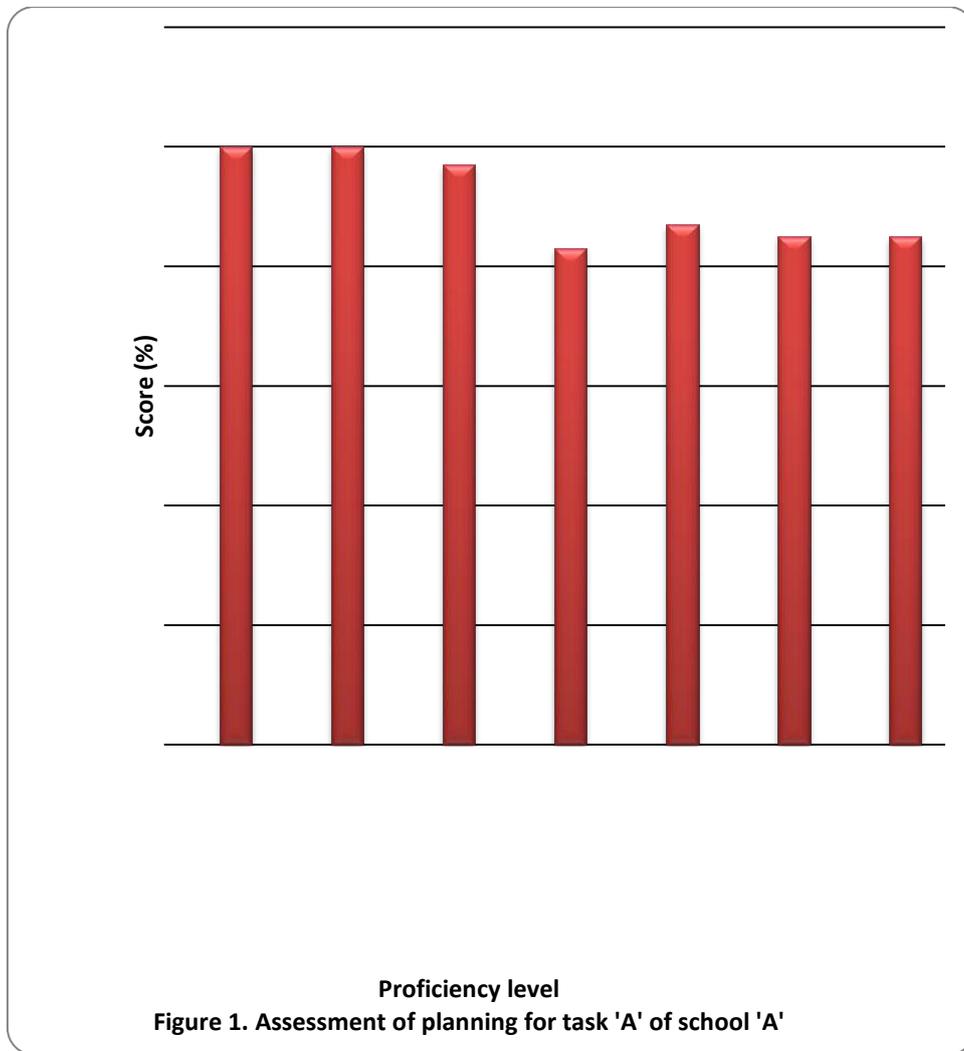
Proficiency was higher than 75% for all the skills exhibited for task 'A' of school B (Figure 2) with the exception of "distil filtrate" and "any safety measure". On the whole, the students scored more than 70% across all the proficiency levels with the exception of "any safety measure".
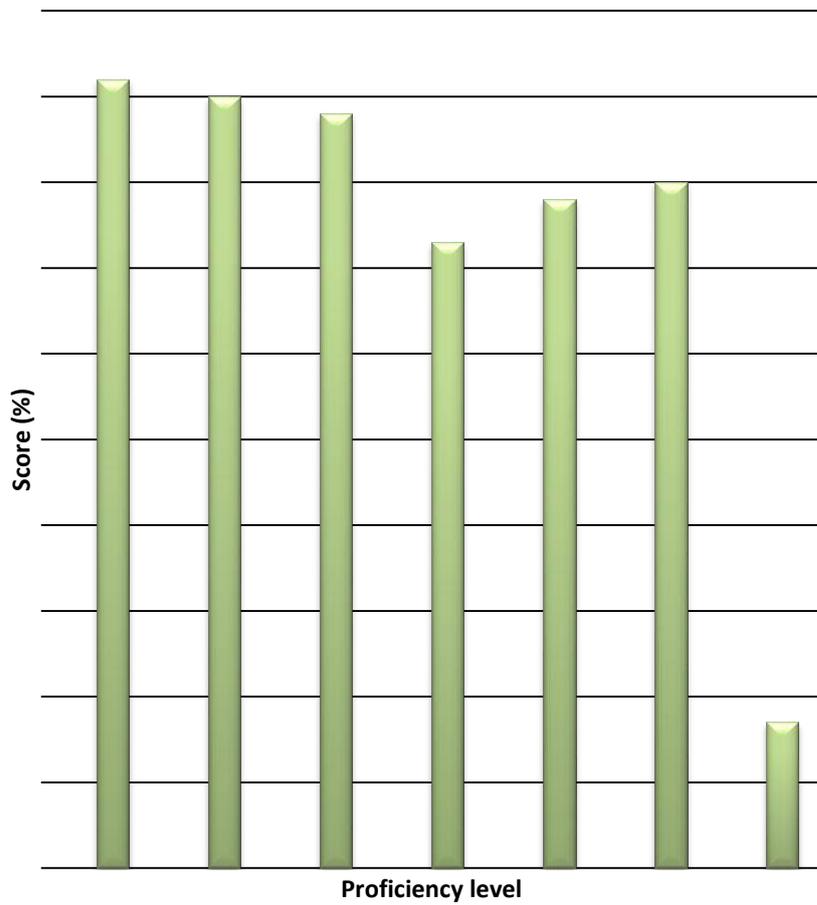
**Figure 2. Assessment of planning for task 'A' of school 'B'**

Generally, Proficiency was higher than 75% for all the skills exhibited for task 'A' of school C (Figure 3). On the whole, all students scored more than 70% across all the proficiency levels with the exception of "any safety measure".

Score (%)
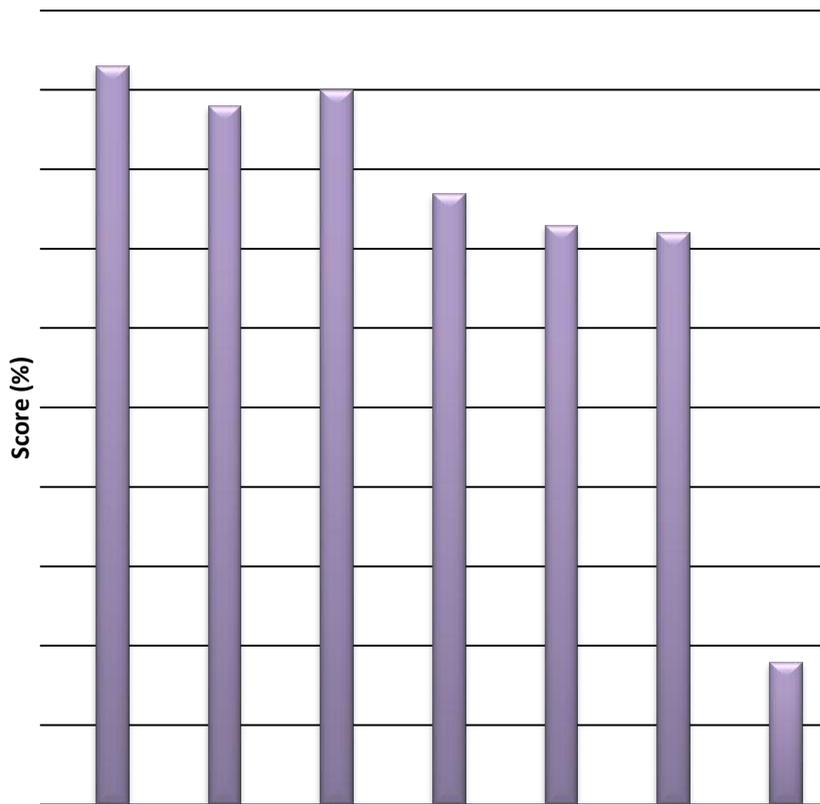
Proficiency level
Figure 3. Assessment of planning for task 'A' of school 'C'

Generally, proficiency was higher than 75% for all the skills exhibited under task A' (Figure 4) with respect to the combined data for the three schools, except "any safety measure".
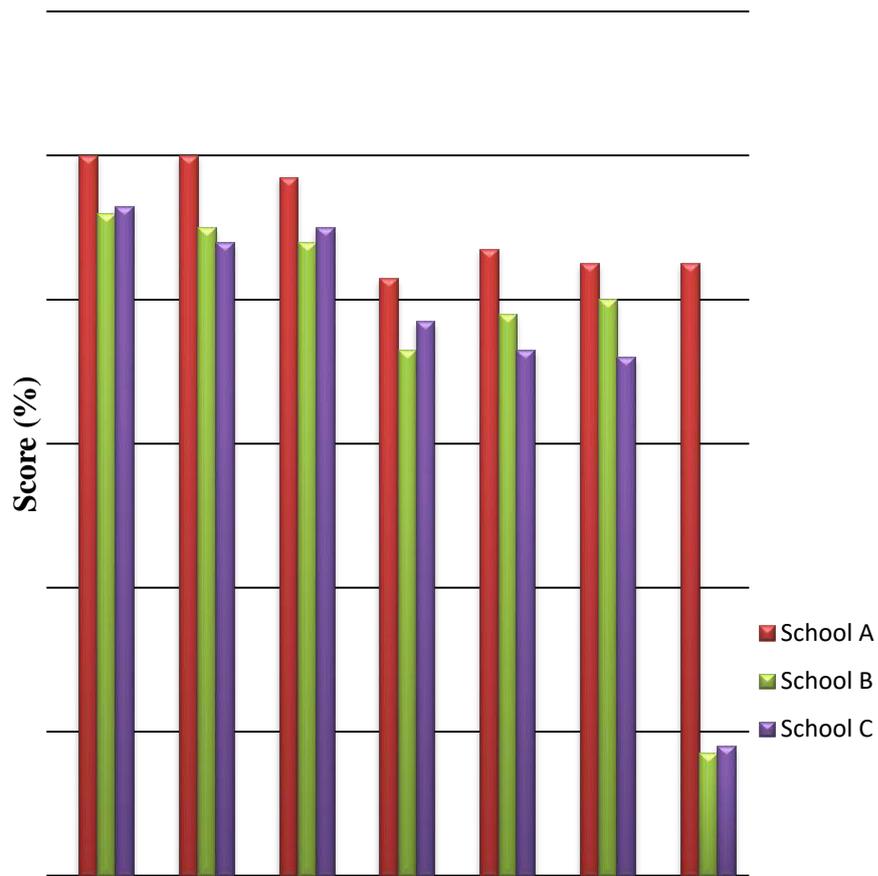
Figure 4. Assessment of planning for task 'A' of the study area

Proficiency level for gender was higher than 75% for all the skills exhibited for task 'A' of school A (Figure 5). A number of students scored more than 80% across all the proficiency levels. On the whole, the males exhibited higher levels of proficiency than females.
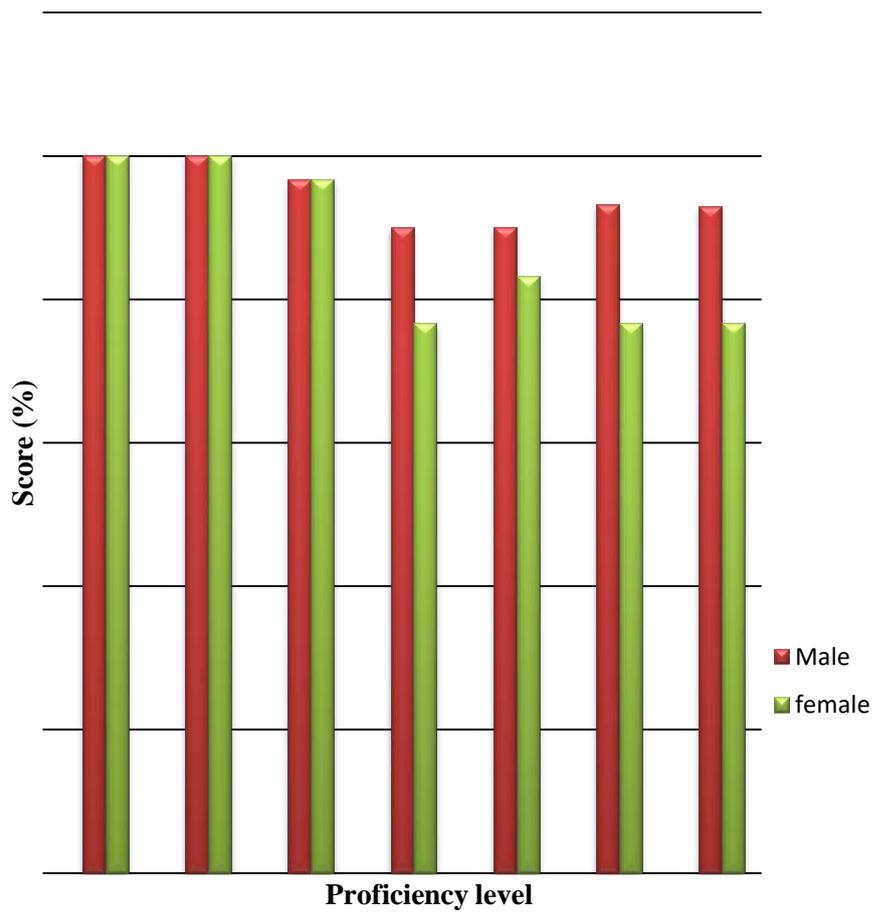
Figure 5. Assessment of gender on plannining skills for task 'A' of school 'A'

Generally, proficiency level of gender was higher than 75% for all the skills exhibited for task 'A' of school B (Figure 6) with the exception of "any safety measures". Almost all the students scored more than 60% across all the proficiency levels. On the whole, the females exhibited higher levels of proficiency than males.
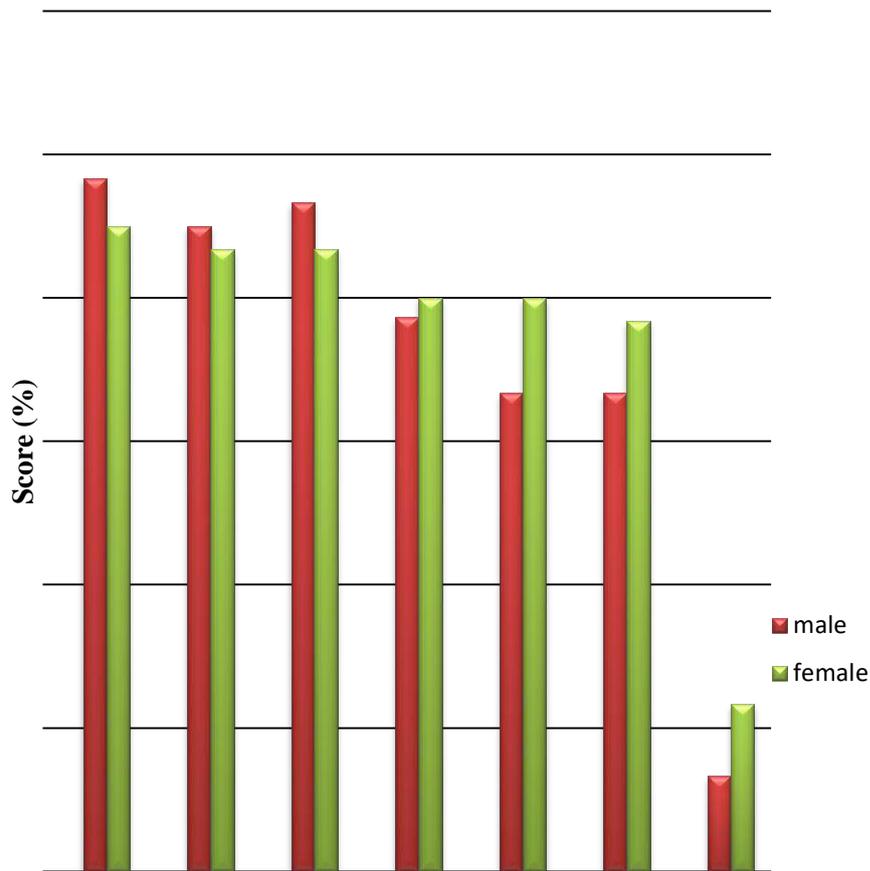
Proficiency level
Figure 6. Assessment of gender on plannining skills for task 'A' of school 'B'

In general, proficiency in terms of gender was higher than 75% for all the skills exhibited for task 'A' of school C (Figure 7) with the exception of "any safety measure". On the whole, the females exhibited higher levels of proficiency than males.
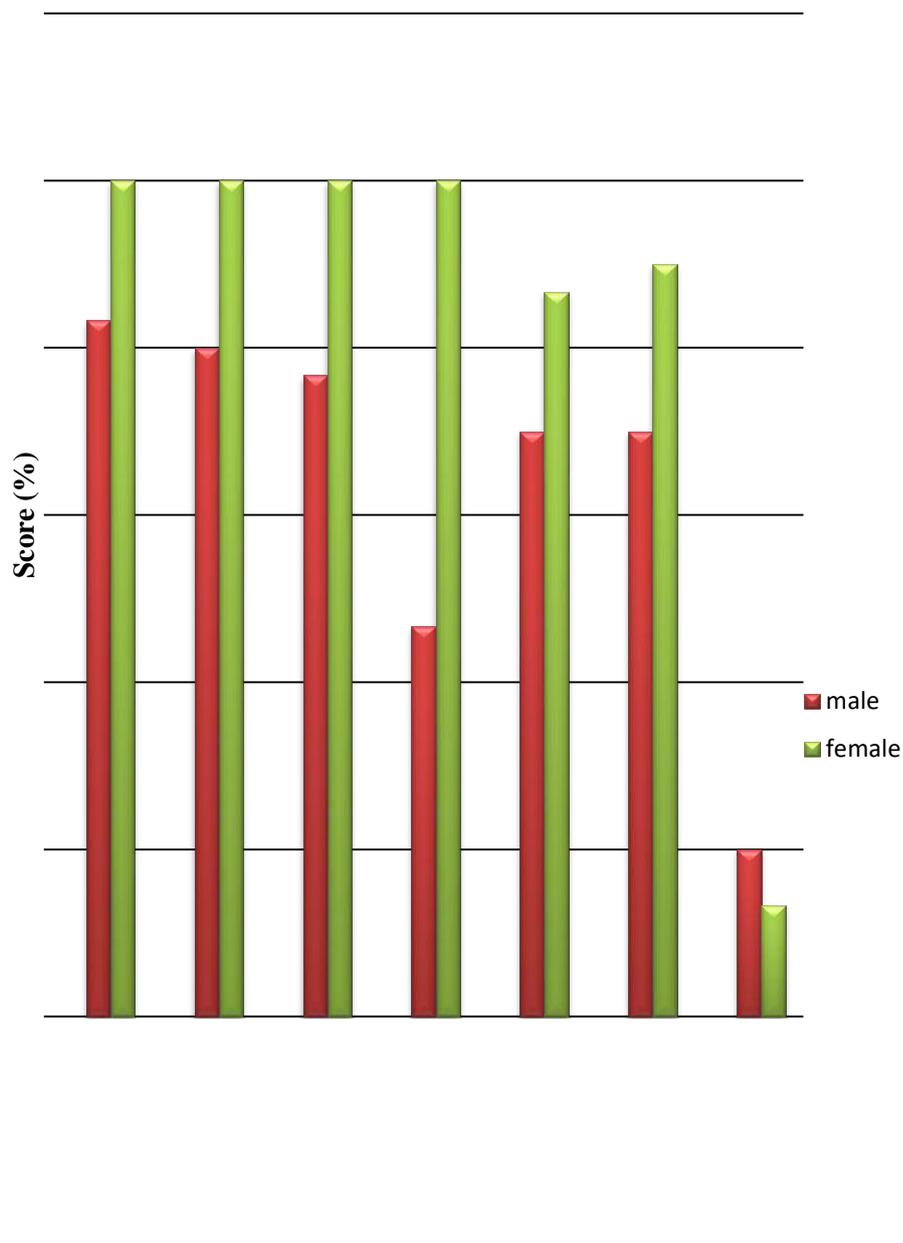
Proficiency level
Figure 7. Assessment of gender on plannining skills for task 'A' of school 'C'

The combined data relating to gender for the three schools shows that in general, proficiency was higher than 75% for all the skills exhibited under task 'A' (Figure 8) with the exception of "any safety measure". On the whole, the female students scored more than 80% across all the skills under the proficiency levels except "any safety measure".
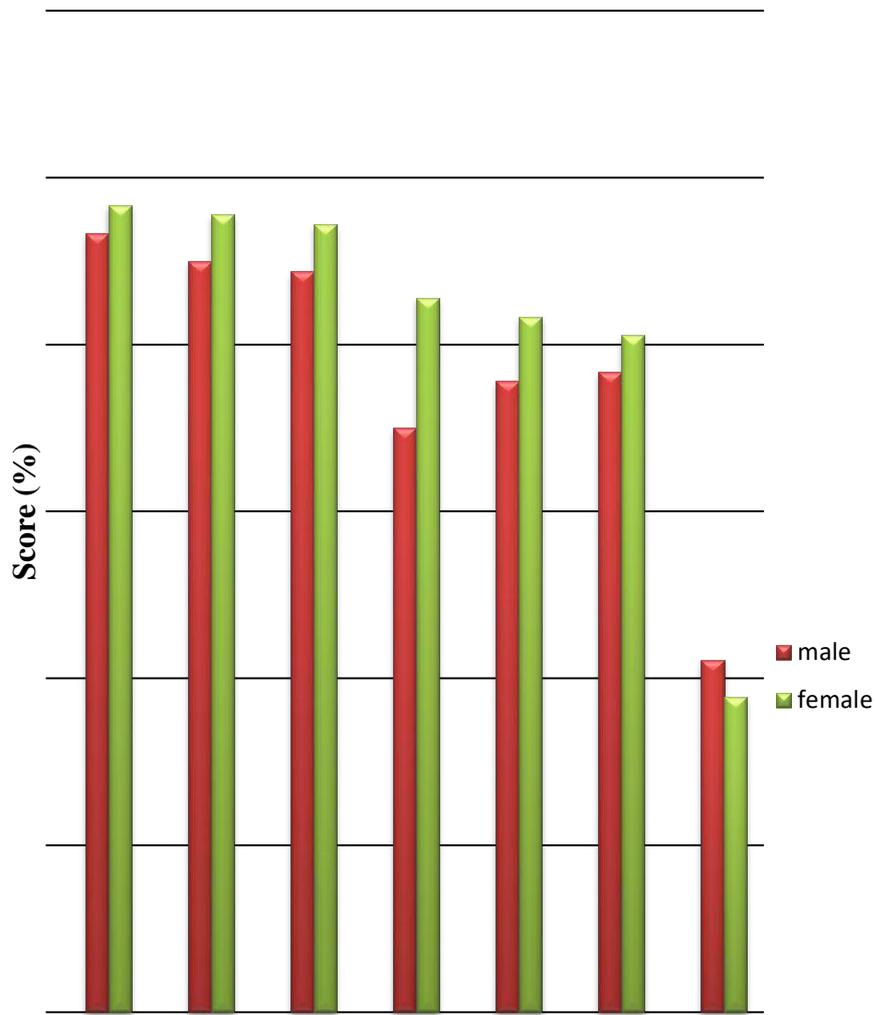
**Figure 8. Assessment of gender on plannining skills for
task 'A' of study area**

(b) Task B: (Appendix B and Appendix E)

Generally, proficiency was higher than 75% for all the skills exhibited for task 'B' of school A (Figure 9) with the exception of "zeroed balance", "mass over volume" and "any safety measure". On the whole, the students scored higher than 80% across all the proficiency levels.
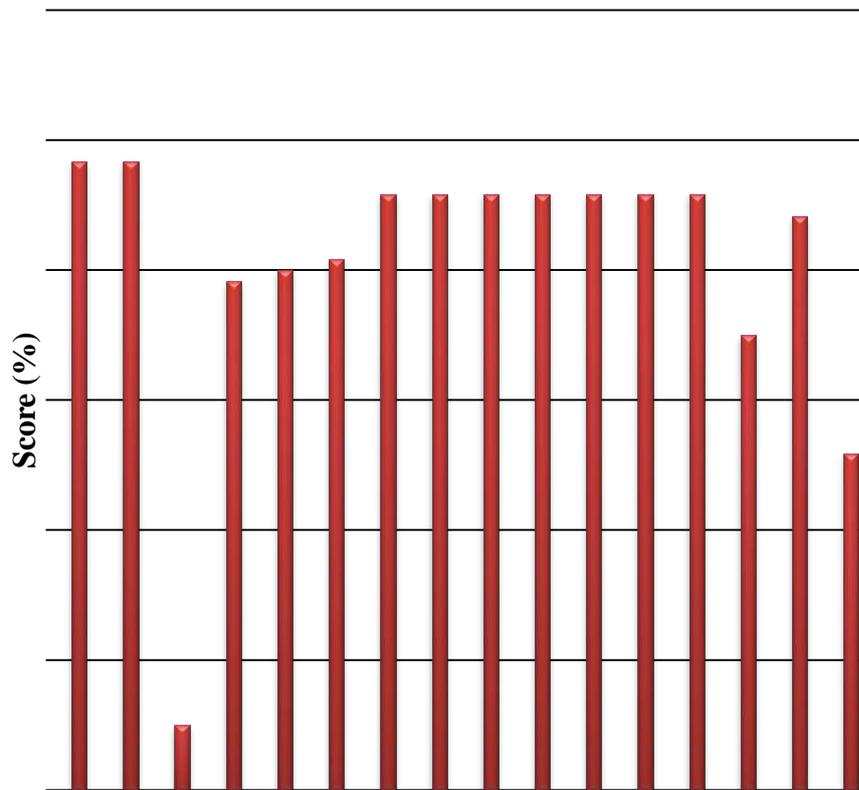
Proficiency level
Figure 9. Assessment of planning for task 'B' of school 'A'

On the whole, proficiency was less than 75% for all the skills exhibited for task 'B' of school B (Figure 10) with the exception of "any plan" and "sequenced order". A number of students scored more than 45% across all the proficiency levels with the exception of "zeroed balance" "mass over volume" and "any safety measure".
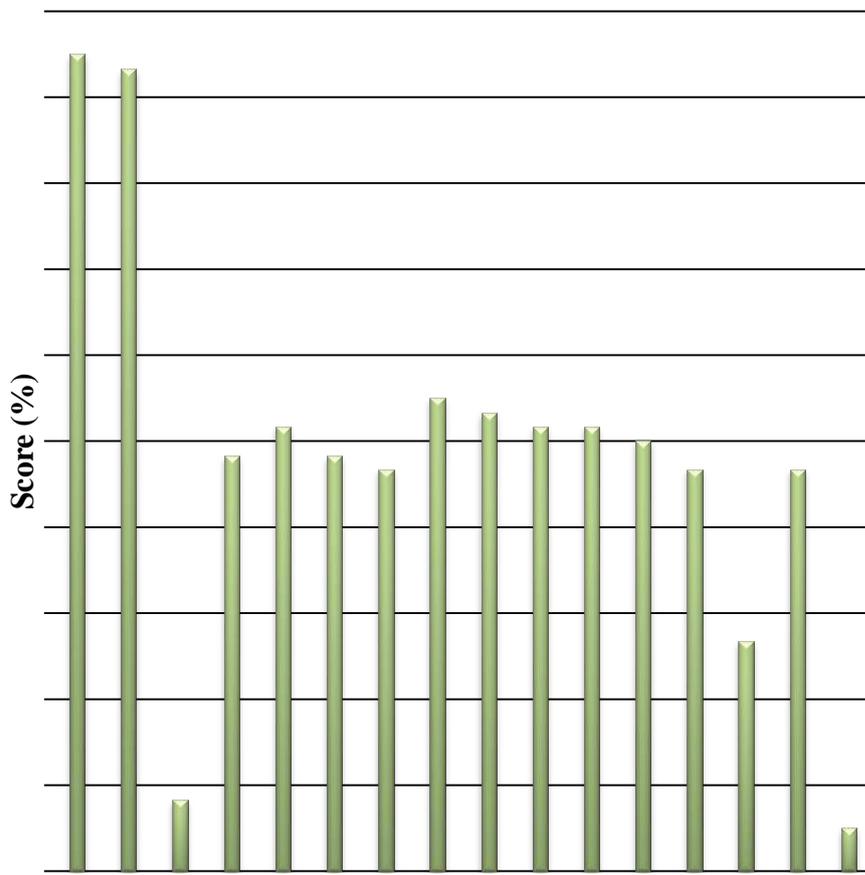
Proficiency level
Figure 10. Assessment of planning for task 'B' of school 'B'

Generally, proficiency was less than 75% for all the skills exhibited for task 'B' under school C (Figure 11) with the exception of "any plan" and "sequenced order". On the whole, the students scored more than 60% across all the proficiency levels.
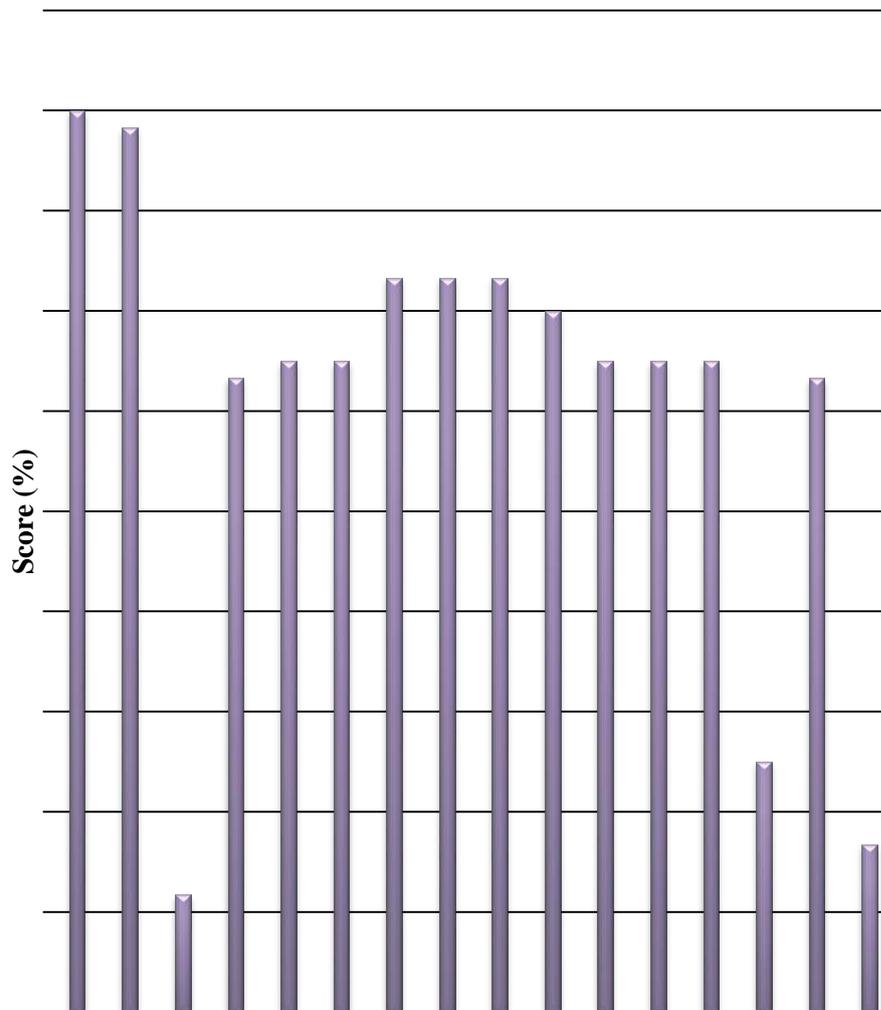
Proficiency level
Figure 11. Assessment of planning for task 'B' of school 'C'

By and large, proficiency was less than 75% for all the skills exhibited under task 'C' with the exception of "any plan" and "sequenced order" (Figure 12) with respect to the combined data of the three schools. However, proficiency higher than 75% for all the skills exhibited for School A.
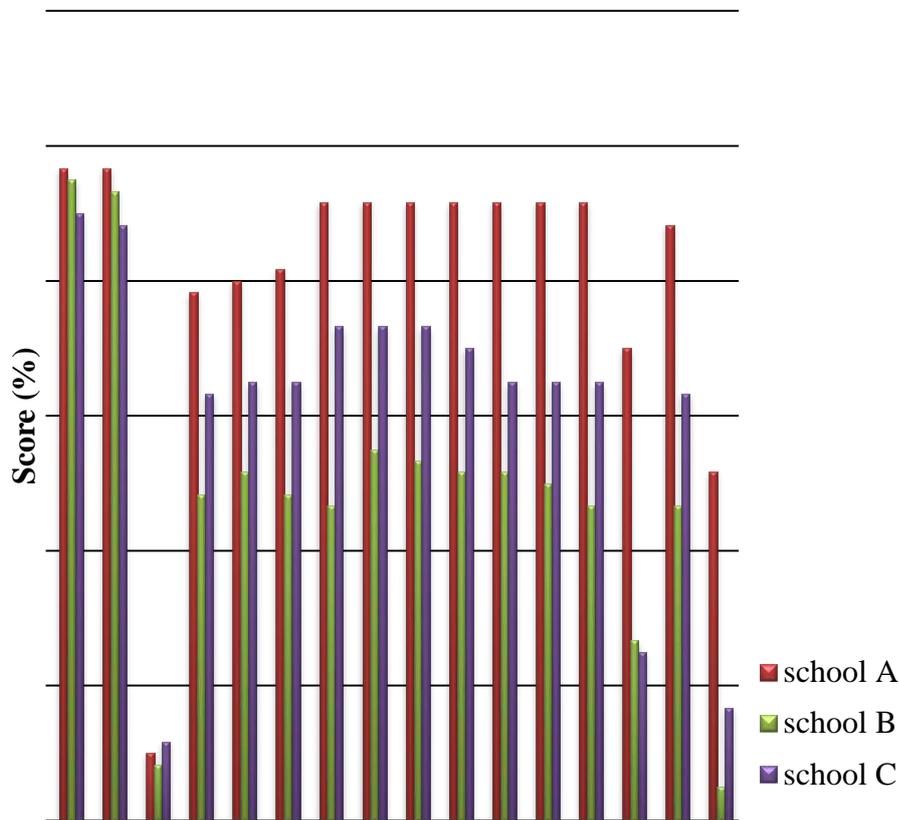
Proficiency level
**Figure 12. Assessment of planning for task 'B' of the
study area**

Generally, proficiency level relating to gender was higher than 75% for all the skills exhibited for task 'B' under school A with the exception of "zeroed balance", and "safety measure" for both sexes, and "gold crown on balance" and "measure mass" for females (Figure 13). On the whole, the male students scored more than 80% across most of the proficiency levels.
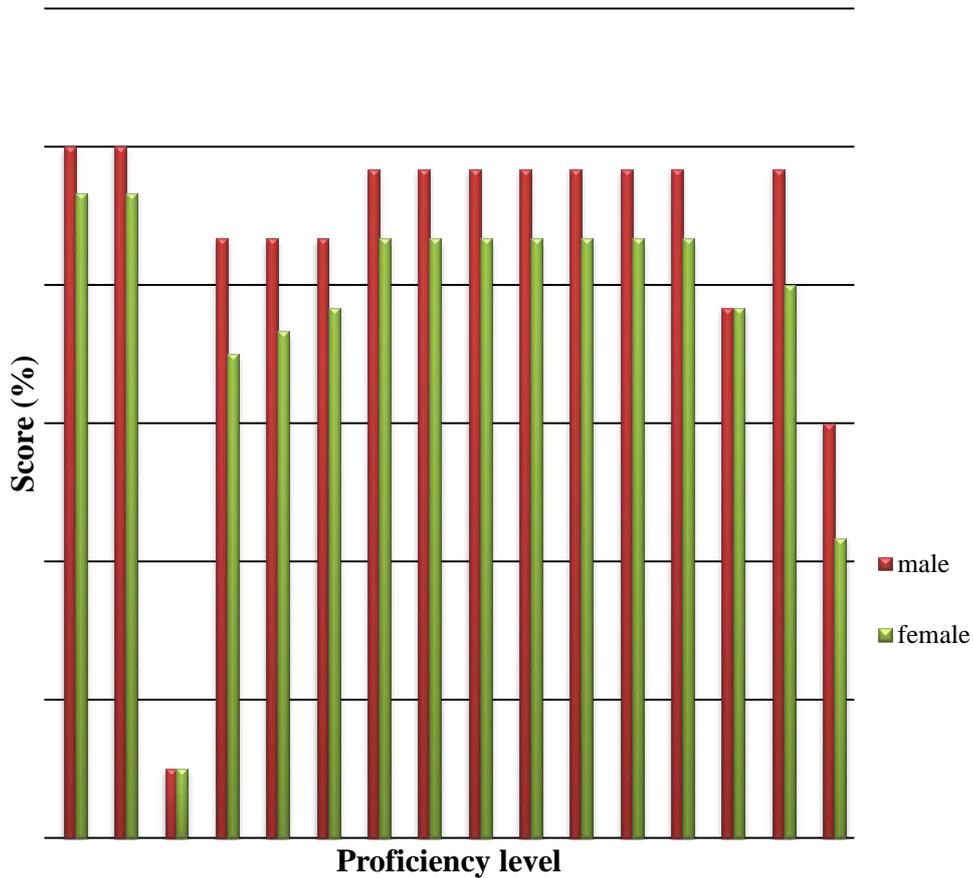
Figure 13. Assessment of gender on plannining skills for task 'B' of school 'A'

On the whole, proficiency level in terms of gender was less than 75% for all the skills exhibited for task 'B' of school B (Figure 14) with the exception of "any plan" and "sequenced order". In general, the male students exhibited higher levels of proficiency than their female counterparts.
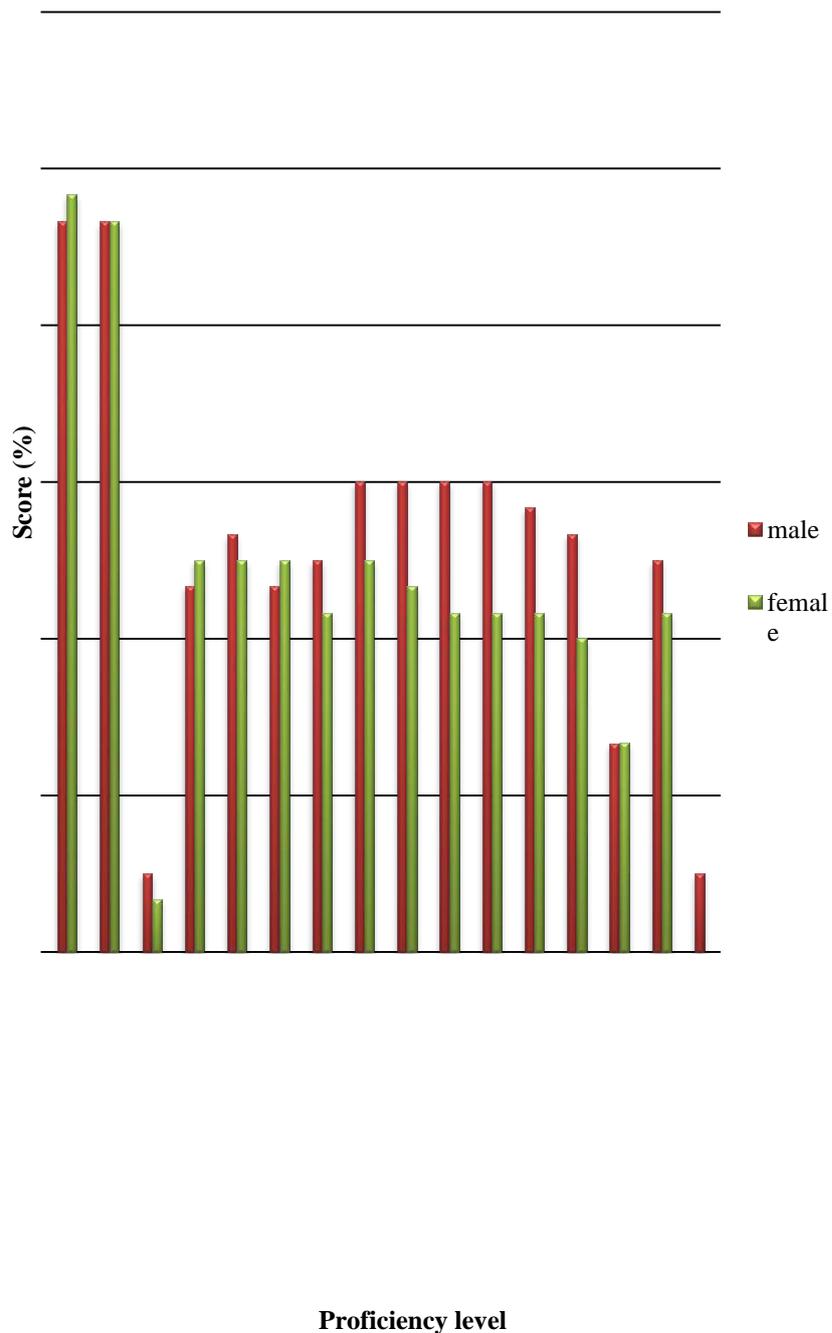
Proficiency level
Figure14. Assessment of gender on plannining skills for task 'B' of school 'B'

In general, proficiency level relating to paired gender was less than 75% for all the skills exhibited for task 'B' under school C (Figure 15) with the exception of "any plan" and "sequenced order". On the whole, the female students exhibited much higher levels of proficiency than their male counterparts.
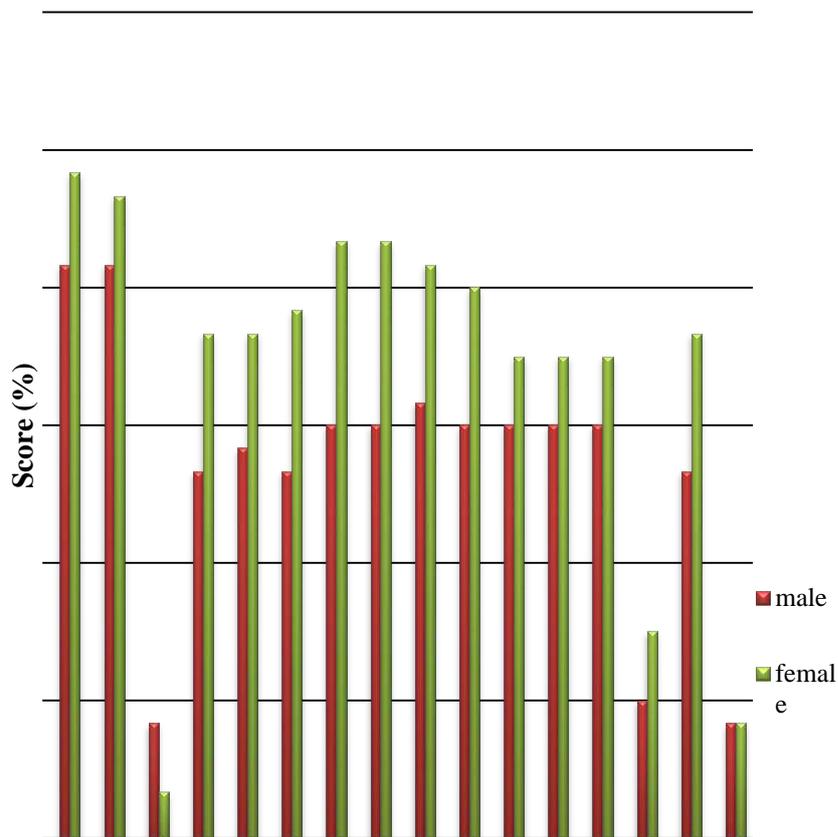
Proficiency level
Figure 15. Assessment of gender on plannining skills for task 'B' of school 'C'

On the whole, proficiency level in terms of gender was less than 75% for all the skills exhibited for task 'B' (Figure 16) with regards to the combined schools of the study area with the with the exception of "any plan" and "sequenced order" Generally, the male students exhibited more skills than their female counterparts.
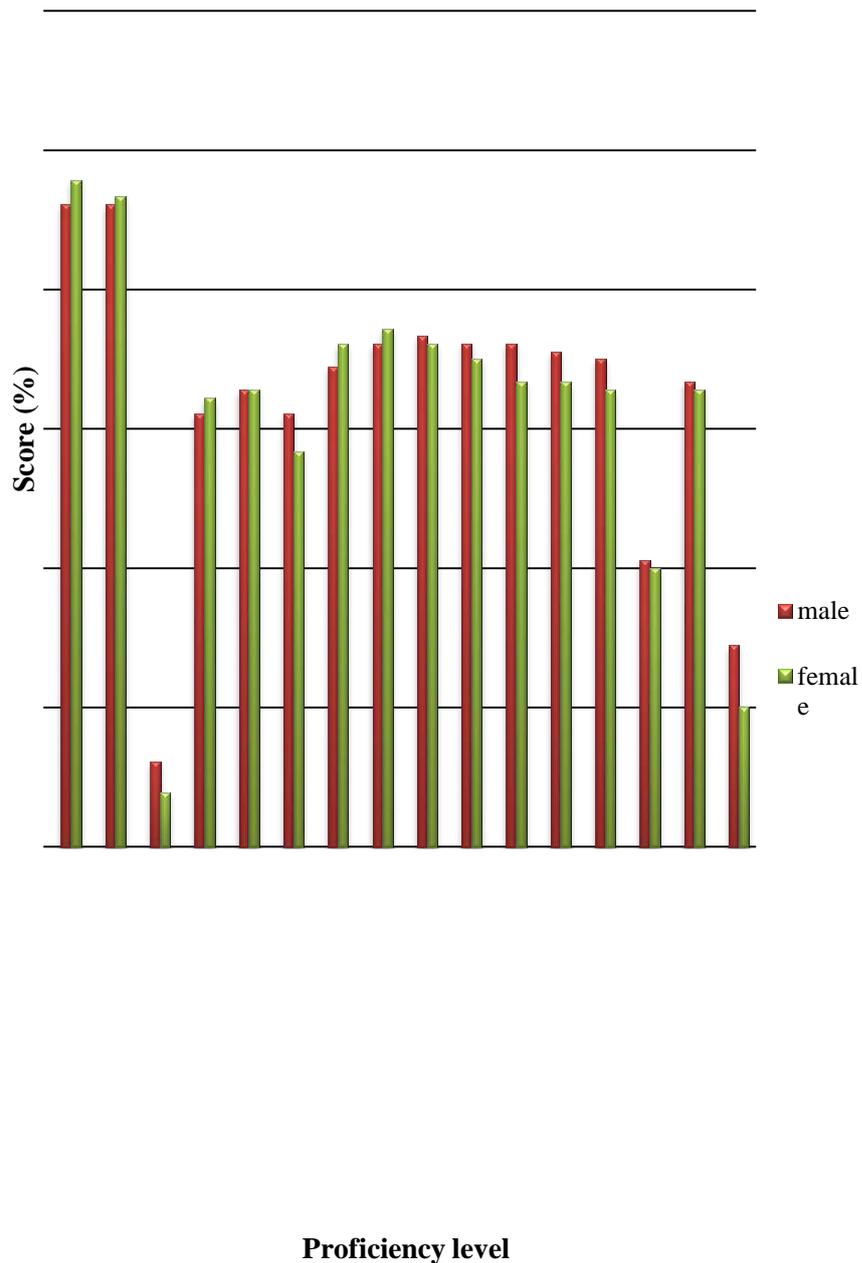
Figure 16. Assessment of gender on plannining skills for task 'B' of the study area

(c) Task C: (Appendix C and Appendix F)

In general, proficiency was higher than 75% for all the skills exhibited for task 'C' of school A (Figure 17) with the exception of "any safety measure". On the whole, the students scored more than 80% across all the proficiency levels.
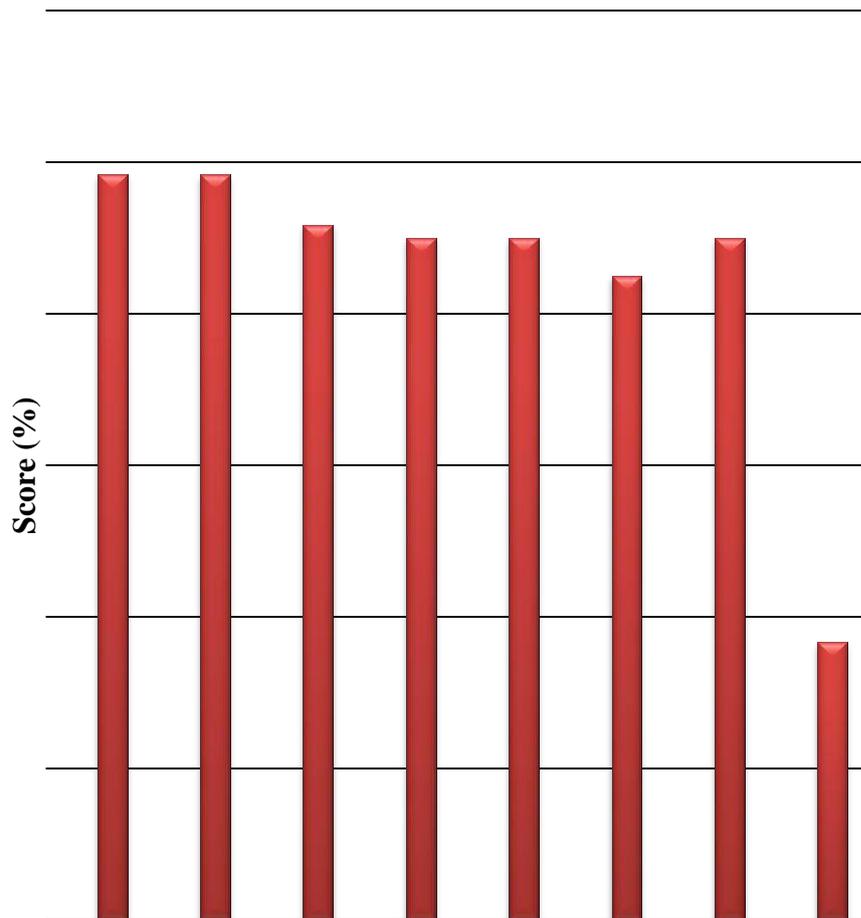
Proficiency level
Figure 17. Assessment of plannining skills for task 'C' of school 'A'

Generally, proficiency was higher than 75% for all the skills exhibited for task 'C' under school B with the exception of "fill beaker", "place yam in beaker", "workable plan" and "safety measure" (Figure 18). On the whole, the students scored more than 60% across all the proficiency levels with the exception of "any safety measure".

Score (%)

Proficiency level
Figure 18. Assessment of plannining skills for task 'C' of school 'B'
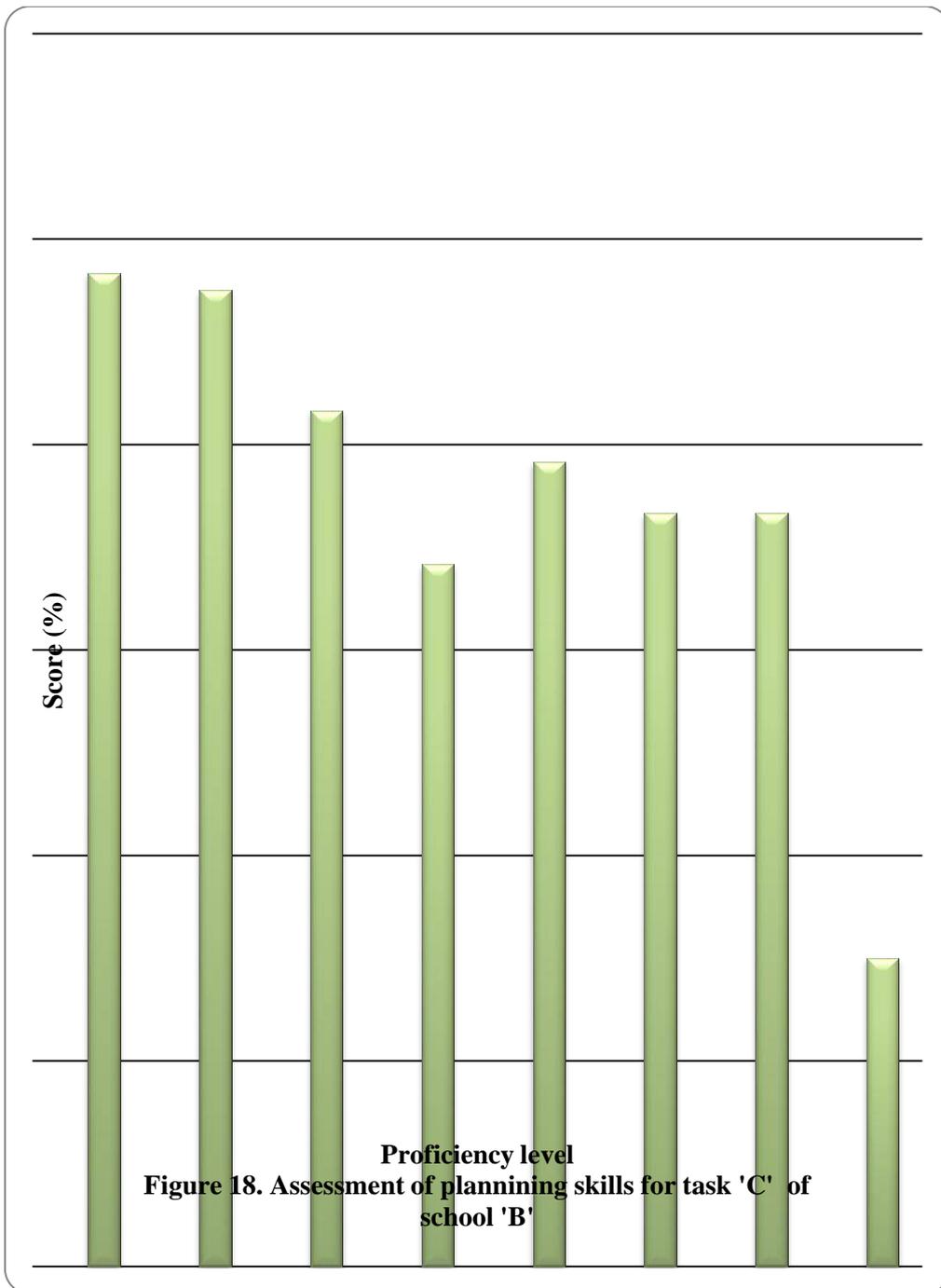
In general, proficiency was higher than 75% for all the skills exhibited for task 'C' in school C (Figure 19) with the exception of "any safety measure". On the whole, the students scored more than 70% across all the proficiency levels.

Score (%)

Proficiency level
Figure 19. Assessment of plannining skills for task 'C' of school 'C'
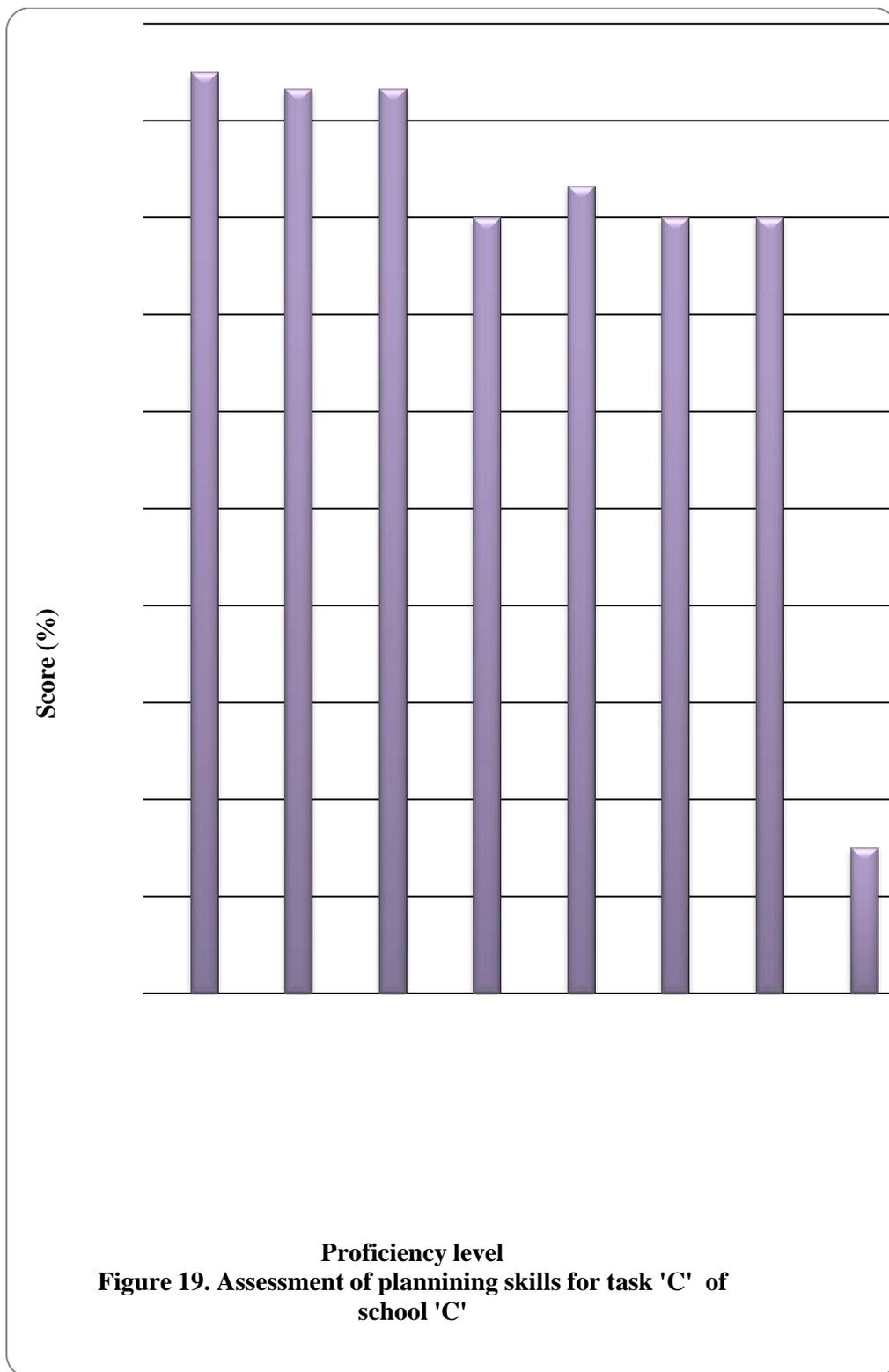
The combined data of the three schools shows that in general, proficiency was higher than 75% for all the skills exhibited under task 'C' with the exception of "safety measure" for all the three schools and "fill beaker", "place yam in beaker", "workable plan" for school B (Figure 20). On the whole, the students scored more than 65% across all the skills under the proficiency levels except "any safety measure".
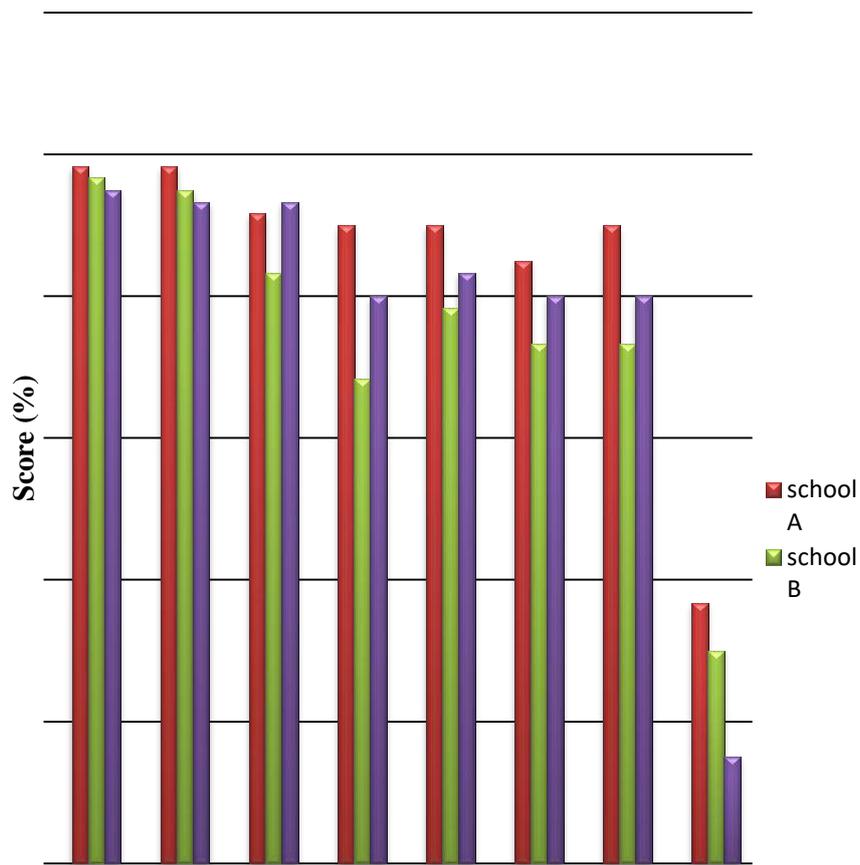
Proficiency level
Figure 20. Assessment of planning for task 'C' of the study area

In general, proficiency level relating to gender was higher than 75% for all the skills exhibited for task 'C' under school A (Figure 21) with the exception of "any safety measure". Female students exhibited higher levels of proficiency than their male counterparts.
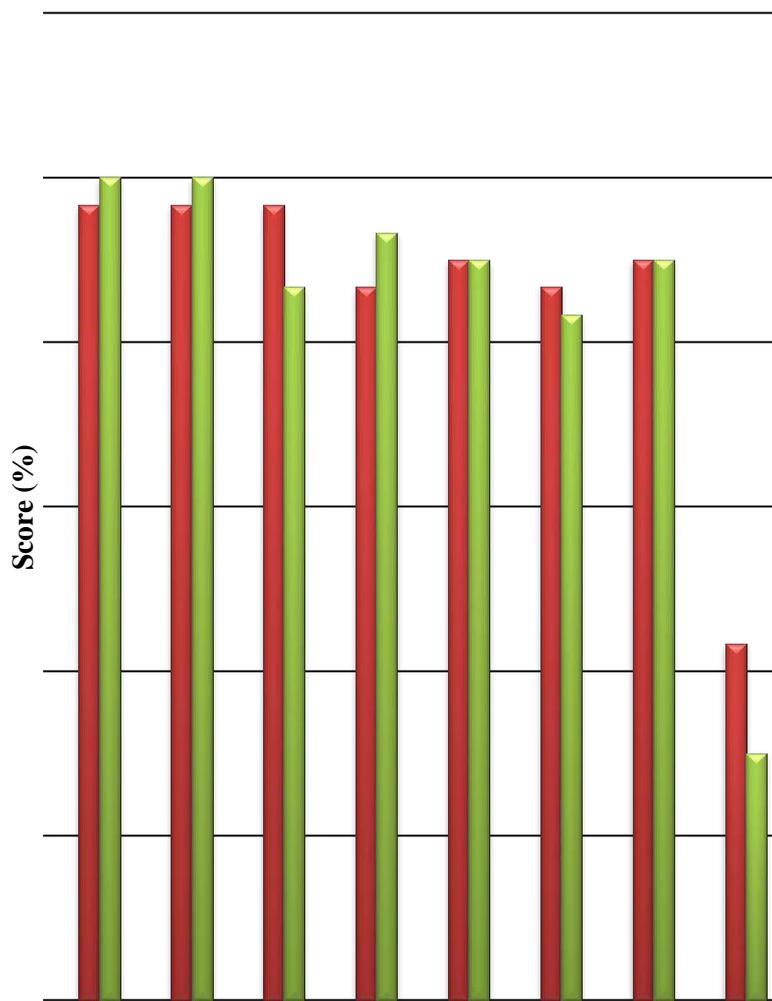
Figure 21. Assessment of gender on plannining skills for task 'C' of school 'A'

Generally, proficiency level in terms of gender was higher than 75% for all the skills exhibited for task 'C' of school B with the exception of "fill beaker" and "safety measure" for both sexes and "peeled and scoped yam", "placed yam in beaker" and "workable plan" for the males (Figure 22). Females exhibited higher levels of proficiency than males.
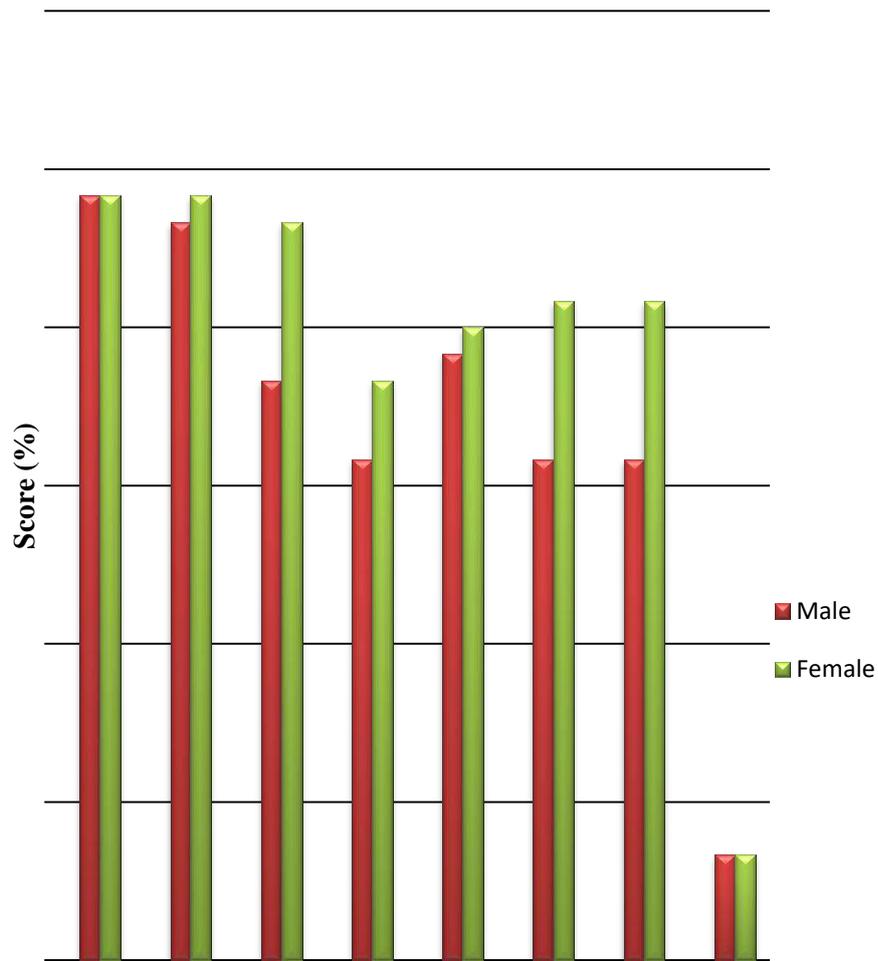
Proficiency level
**Figure 22. Assessment of gender on plannining skills for task
'C' of school 'B'**

Generally, proficiency level relating to gender was higher than 75% for all the skills exhibited for task 'C' under school C with the exception of "safety measure" for both sexes and "fill beaker", fill yam with salt", place yam in beaker" and "workable plan" for males (Figure 23). Females exhibited higher levels of proficiency than males.
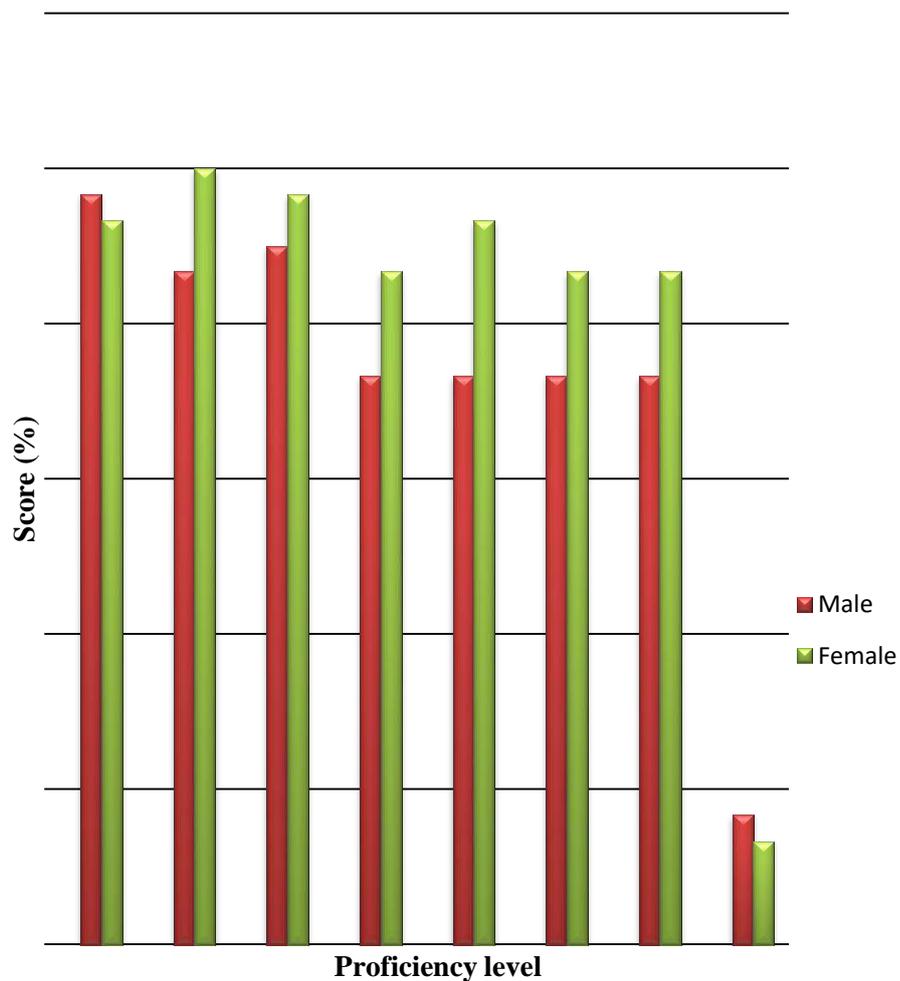
Figure 23. Assessment of gender on plannining skills for task 'C' of school 'C'

Generally, proficiency in terms of gender was higher than 75% for all the skills exhibited under task 'C' with the exception of "safety measure" for both sexes and "fill beaker" and "place yam in beaker" for males (Figure 24) with respect to the combined data of the three schools. On the whole, the female students scored more than 80% across all the proficiency levels compared to their male counterparts.
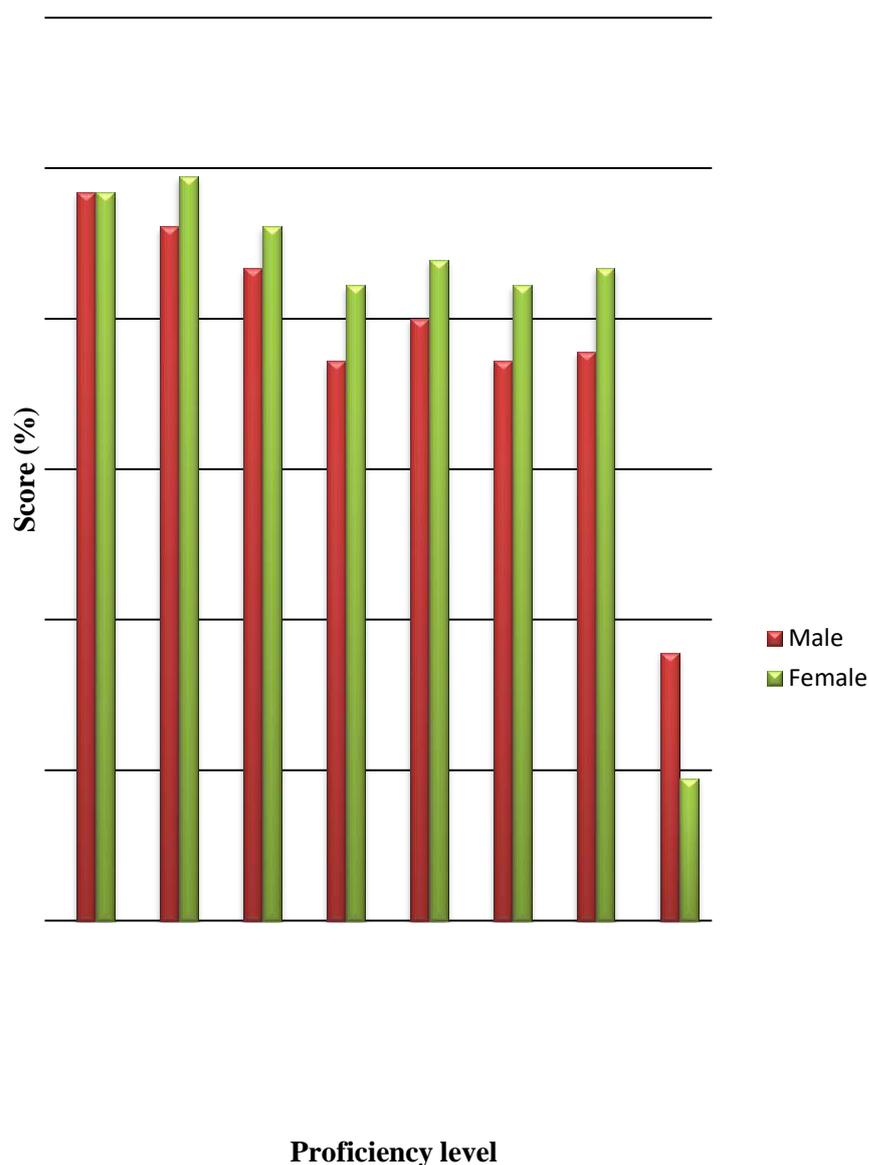
Proficiency level
Figure 24. Assessment of gender on plannining skills for task 'C' of the study area

Most students outlined the steps to execute the tasks (Figures 1, 2, 3, 4, 9, 10, 11, 12, 17, 18, 19 and 20) but could not state safety measures. This negates the statement that a great number of students strategies for planning experiments, are often unsystematic (Kuhn & Phelps, 1982; Hammann et al., 2008). The claim that student conclusions from experimental data are often invalid and driven by confirmation bias (Chinn and Brewer, 1986; Hammann et al., 2008) was not supported by the current study. Thus, the results of this study revealed that students possessed a scientific understanding of process skills involved in conducting an experiment, which is also revealed in the performance test by the great number of students who exhibited planning skills in their tasks.

The performance test gave far greater freedom for students to design individual tasks. One of the positive findings of this study is that students wrote valid procedure for their own planned tasks, though they were not able to state valid safety measure for planned activity. Students' performance on task C was most encouraging compared to the rest of the tasks (Figure 20). This could be due to the fact that students were most probably familiar with the materials provided (yam and salt) and the task was more related to everyday real life situation. Task B was challenging as it involved using a lot of the processing skills in planning the task (Figure 12). However, a research conducted by Liu (2000) indicates that, students do not need to

acquire a vast amount of information, but rather the ability to think, and organise information for specific purposes. This is reflected in the current study whereby the mean score (10.28, out of 16) was high.

The study has revealed that on the whole, the proficiency levels of most of the students were higher than 75% on tasks A and C of the planning skill for all the schools. But a close look at the students answer booklet revealed that only a few students (30.6%) were able to state safety measures for all the three tasks. However the proficiency levels exhibited by school A on task B were higher than 75% whiles that of schools B and C were less than 75%.

Research Question Two: Which gender shows more proficiency in planning skills?
Specific Details

(a) Task A

Tables 1, 3, and 5 display the total number of scores exhibited by the students of schools A, B, and C respectively on task A. In Table 1, most of the students displayed more than two proficient skills in planning the task. Specifically, the students exhibited six and seven of the skills. In Table 3, few students could not display more than two of the skills. However, most of the students exhibited four of the skills whiles few displayed all the skills. On the other hand, in Table 5, few of the students could not display any skill. Nevertheless, most of the students exhibited six of the skills whiles few displayed all the skills.

Table 1: Distribution of Total Score by Gender on Task A of School A

Total scores	Male Number of students (n = 30)	Female Number of students (n = 30)	Total number of students (n = 60)
0	0	0	0
1	0	0	0
2	1	0	1
3	0	2	2
4	2	3	5
5	3	3	6
6	13	11	24
7	11	11	22

Table 2 depicts a contrast between the mean performance of males and females (6.00 ± 1.15 and 5.87 ± 1.22 respectively) for students of school A on task A. However, the mean performance of the total number of students was 5.93± 1.18.

Table 2: Distribution of p and t Values by Gender of Students of School A on Tasks A

Task A School A	Mean ± SD	p value	T	
Male	6.00 ± 1.15	0.665	0.436	
Female	5.87 ± 1.22			
Total	5.93± 1.18			

From the t-value calculated there is no statistically significant difference between performance of male and female (t(58) = 0.436, p = 0.665).

Table 3: Distribution of Total Score by Gender on Task A of School B

Total scores	Male Number of students (n= 30)	Female Number of students (n = 30)	Total number of students (n = 60)
0	1	3	4

1	1	0	1
2	1	1	2
3	5	2	7
4	3	3	53
5	1	1	2
6	16	17	33
7	2	6	8

As shown in Table 4, the mean performance of males and females was similar (4.93± 1.80 and 5.13± 2.16 respectively) for students of school B on task A. Nonetheless, the mean performance of the total number of students was 5.03± 1.97.

Table 4: Distribution of p and t Values by Gender of Students of School B   on Tasks A

Task A School B	Mean ± SD	p value	t
Male	4.93± 1.80	0.698	0.390
Female	5.13± 2.16		
Total	5.03± 1.97		

From the t-value calculated there is no statistically significant difference between performance male and female (t(58) = 0.390, p = 0.698).

Table 5: Distribution of Total Score by Gender on Task A of School C

Total scores	Male Number of students (n = 30)	Female Number of students (n = 30)	Total number of students (n = 60)
0	4	0	4
1	1	0	1
2	0	0	0
3	2	0	2
4	1	2	3
5	6	5	11
6	14	20	34
7	2	3	5

Table 6 depicts a contrast between the mean performance of males and females (4.63 ± 2.22 and 5.60 ± 1.28 respectively) for students of school C on task A. However, the mean performance of the total number of students was 5.12 ± 1.86.

Table 6: Distribution of p and t Values by Gender of Students of School C on Tasks A

Task A School C	Mean ± SD	p value	T	
Male	4.63 ± 2.22	0.043	2.068	
Female	5.60 ± 1.28			
Total	5.12 ± 1.86			

From the t-value calculated there is a statistically significant difference between performance of males and females (t(58) = 2.068, p = 0.043).

Table 7 depicts a similarity between the mean performance of males and females (5.19 ± 1.85 and 5.53 ± 1.62 respectively) for students of schools for the study area on task A. However, the mean performance of the total number of students was 5.36 ± 1.75.

Table 7: Distribution of p and t Values by Gender of Students of Schools for the Study Area on Tasks A

Task A Schools for study area	Mean ± SD	p value	t	
Male	5.19 ± 1.85	0.186	1.326	
Female	5.53 ± 1.62			
Total	5.36 ± 1.75			

From the t-value calculated there is no statistically significant difference between performance of males and females (t(178) = 1.326, p = 0.186).

(b) Task B

Tables 8, 10, and 12 show the total number of scores exhibited by students of schools A, B, and C respectively on task B. In Table 8, few of the students displayed less than two skills in planning the task. Few as well exhibited between 9 and 13 of the skills, whiles most of the students exhibited 14, 15 and 16 of the skills. Table 10 shows that a few students could not display any of the skills. However, most of the students exhibited between 2 and 13 of the skills whiles few displayed between 14 and 15 of the skills. None of the students could exhibit all the skills. On the other hand, in Table 12, few of the students could not display any skill. Nevertheless, few of the students exhibited between 2 and 12 of the skills whiles a considerable number displayed between 13 and 15 all the skills but none of them could display all the skills.

Table 8: Distribution of Total Score by Gender on Task B of School A

Total scores	Male Number of students (n = 30)	Female Number of students (n = 30)	Total number of students (n = 60)
0	1	1	2
1	0	1	1
2	1	2	3
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
9	1	0	1
10	2	0	2
11	1	3	4
12	0	1	1
13	3	3	6
14	4	8	12
15	13	8	21
16	4	3	7

As shown in Table 9, the mean performance of males and females was similar (13.13 ± 3.79 and 12.37 ± 4.65 respectively) of students of school A on task B. Nonetheless, the mean performance of the total number students was 12.75 ± 4.22.

Table 9: Distribution of p and t Values by Gender of Students of School A on Tasks B

Task B School A	Mean ± SD	p value	t	
Male	13.13 ± 3.79	0.486	0.700	
Female	12.37 ± 4.65			
Total	12.75 ± 4.22			

From the t-value calculated there is no statistically significant difference between performance of male and female (t(58) = 0.700, p = 0.486).

Table 10: Distribution of Total Score by Gender on Task B of School B

Total score	Male Number of students (n = 30)	Female Number of students (n = 30)	Total number of students (n = 60)
0	1	6	7
1	0	0	0

Table 10 (continued)

Total score	Male Number of students (n = 30)	Female Number of students (n = 30)	Total number of students (n = 60)
2	6	6	12
3	2	1	3
4	1	1	2
5	0	0	0
6	1	1	2
7	0	0	0
8	1	0	1
9	0	2	2
10	1	0	1
11	0	1	1
12	0	0	0
13	6	3	9
14	8	7	15
15	3	2	5
16	0	0	0

Table 11 depicts a contrast between the mean performance of males and females (9.37 ± 5.52 and 7.37 ± 6.07 respectively) of students of school B on task B. However, the mean performance of the total number students was 8.37 ± 5.84.

Table 11: Distribution of p and t Values by Gender of Students of School B on Tasks B

Task B School B	Mean ± SD	p value	t	
Male	9.37 ± 5.52	0.187	1.335	
Female	7.37 ± 6.07			
Total	8.37 ± 5.84			

From the t-value calculated there is no statistically significant difference between performance of males and females (t(58) = 1.335, p = 0.187).

Table 12: Distribution of Total Score by Gender on Task B of School C

Total score	Male Number of students (n = 30)	Female Number of students (n = 30)	Total number of students (n = 60)
0	5	1	6
1	0	0	0

Table 12 (continued)

Total score	Male	Female	Total number of students (n = 60)
	Number of students (n = 30)	Number of students (n = 30)	
2	6	2	8
3	0	0	0
4	0	1	1
5	0	1	1
6	1	0	1
7	0	1	1
8	0	0	0
9	1	2	3
10	1	3	4
11	1	0	1
12	1	0	1
13	4	9	13
14	8	6	14
15	2	4	6
16	0	0	0

Table 13 shows a contrast between the mean performance of males and females (8.47 ± 6.02 and 10.97 ± 4.35 respectively) of students of school C on task B. However, the mean performance of the total number students was 9.72 ± 5.36.

Table 13: Distribution of p and t Values by Gender of Students of School C on Tasks B

Task B School C	Mean ± SD	p value	t	
Male	8.47 ± 6.02	0.071	1.844	
Female	10.97 ± 4.35			
Total	9.72 ± 5.36			

From the t-value calculated there is no statistically significant difference between performance of male and female (t(58) = 1.844, p = 0.071).

Table 14 describes a comparable mean performance of males and females (10.32 ± 5.52 and 10.23 ± 5.45 respectively) of students of schools of the study area on task B. Nonetheless, the mean performance of the total number students was 10.28 ± 5.47.

Table 14: Distribution of p and t Values by Gender of Students of Schools for the Study Area on Tasks B

Task B Schools for study area	Mean ± SD	p value	t	
Male	10.32 ± 5.52	0.941	0.109	
Female	10.23 ± 5.45			

Total	10.28 ± 5.47			

From the t-value calculated there is no statistically significant difference between the performance of males and females (t(178) = 0.109, p = 0.941).

(c) Task C

Tables 15, 17, and 19 show the total number of scores exhibited by students of schools A, B, and C respectively on task C. In Table 15, only 1 of the students could not display any skill in planning the task. Also few of the students exhibited between 2 and 6 of the skills, whiles most of them exhibited between 7 and 8 of the skills. In Table 17, one student could not display any of the skills. However, most of the students exhibited between 2 and 6 of the skills whiles the rest displayed between 7 and 8 of the skills. On the other hand, in Table 19, 1 student could not display any skill. Nonetheless, few of the students exhibited between 1 and 6 of the skills whereas most of them displayed between 7 and 8 the skills.

Table 15: Distribution of Total Score by Gender on Task C of School A

Total score	Male Number of students (n = 30)	Female Number of students (n = 30)	Total number of students (n = 60)
0	1	0	1
1	0	0	0
2	0	1	1
3	1	1	2
4	1	0	1
5	1	1	2
6	1	6	7
7	12	12	24
8	13	9	22

Table 16 describes a comparable mean performance of males and females (6.83 ± 1.76 and 6.77 ± 1.43 respectively) of students of school A on task C. Nonetheless, the mean performance of the total number students was 6.80 ± 1.59.

Table 16: Distribution of p and t Values by Gender of Students of School A on Tasks C

Task C School A	Mean ± SD	p value	t	
Male	6.83 ± 1.76	0.873	0.161	
Female	6.77 ± 1.43			
Total	6.80 ± 1.59			

From the t-value calculated there is no statistically significant difference between the performance of males and females (t(58) = 0.161, p = 0.873).

Table 17: Distribution of Total Score by Gender on Task C of School B

Total score	Male Number of students (n = 30)	Female Number of students (n = 30)	Total number of students (n = 60)
0	1	0	1
1	1	1	2
2	3	3	6
3	2	0	2

4	3	1	4
5	1	1	2
6	4	4	8
7	12	16	28
8	3	4	7

Table 18 shows a contrast between the mean performance of males and females (5.43 ± 2.22 and 6.20 ± 1.85 respectively) of students of school B on task C. However, the mean performance of the total number students was 5.82 ± 2.06.

Table 18: Distribution of p and t Values by Gender of Students of School B on Tasks C

Task C School B	Mean ± SD	p value	T	
Male	5.43 ± 2.22	0.152	1.453	
Female	6.20 ± 1.85			
Total	5.82 ± 2.06			

The t-value calculated there is no statistically significant difference between the performance of males and females ($t(58) = 1.453$, $p = 0.152$).

Table 19: Distribution of Total Score by Gender on Task C of School C

Total score	Male	Female	Total number of students (n = 60)
	Number of students (n= 30)	Number of students (n= 30)	
0	1	0	1
1	2	0	2
2	1	0	1
3	3	2	5
Table 19 (continued)			
Total score	Male	Female	Total number of students (n = 60)
	Number of students (n= 30)	Number of students (n= 30)	
4	1	2	3
5	0	0	0
6	1	1	2
7	18	20	38
8	4	4	8

As shown in Table 20, the mean performance of males and females was similar (6.03 ± 2.13 and 6.40 ± 1.79 respectively) for students of school C on task C. Nonetheless, the mean performance of total number of students was 6.22 ± 1.96.

Table 20: Distribution of p and t Values by Gender of Students of School t B on Tasks C

Task C School B	Mean ± SD	p value	t	
Male	6.03 ± 2.13	0.473	0.722	
Female	6.40 ± 1.79			
Total	6.22 ± 1.96			

The t-value calculated there is no statistically significant difference between the performance of males and females (t(58) = 0.722, p = 0.473). As shown in Table 21, the mean performance of males and females was similar (6.10 ± 2.10 and 6.46 ± 1.70 respectively) of students of schools for the study area on task C. Nonetheless, the mean performance of the total number students was 6.28 ± 1.92.

Table 21: Distribution of p and t Values by Gender of Students of Schools for the Study Area on Tasks C

Task C Schools for study area	Mean ± SD	p value	T	
Male	6.10 ± 2.10	0.214	1.248	
Female	6.46 ± 1.70			
Total	6.28 ± 1.92			

The t-value calculated there is no statistically significant difference between performance of male and female (t(178) = 1.248, p = 0.214).

Gender differences were analyzed on each task. For each task, an independent *t*-test was conducted to compare the tasks for which the gender difference was significant at the 5% level of probability. There was no significant difference in the performance of male and females on task A for schools A and B. This finding underpins the study conducted by Pine et al. (2006), who found comparable gender performance on physical science tasks. However, there was a significant difference in the performance of males and females on task A for school C. Females outperformed their male counterparts. A number of studies support this finding (Klein et al., 1997; Pine et al., 2006; Shaw & Nagashima, 2009). On the whole, the mean scores of this study revealed that females score was higher than males on task A for all the three schools in the study area, even though the difference was not statistically significant.

Tables 9 and 11 indicate that, the mean score of the males was much higher than that of the females of school A and B on task B. Nonetheless, the t-test value indicates that the difference in performance of males and females was not significant. Females mean score (Table 13) was higher than that of the males on task B of school C. However, the difference in performance was not statistically significant.

Gender performance across the schools of the study area did not show any significant difference in performance (p > 0.05) as shown in Table 14. The students performed relatively better on the skill of recording values. The current study corroborates the findings of Beaumont-Walters and Soyibo (2001) that students performed relatively better on the skill of recording data. One of the findings was students' inability to zero balance before measuring. This was observed for most of the students in all the schools that were investigated.

Analysis of data for task C shown in Tables 18 and 20 (schools B and C) depicts that the mean scores of females was higher than the males. The converse holds true for school A. However, the t-test value indicates that there were no statistically significance difference between the performance of female and males. However, the t-value on gender for all the schools did not indicate a significant difference between male and female performance. This current study supports the findings of Ssempala (2008) who asserted that there are more similarities than differences between the performances of females and males on performance assessment tasks.

On the whole, students had difficulty in stating safety measure when planning the tasks.

Research Question Three: Which students of the various school types exhibit most proficiency level in planning skills?

(3) ANOVA Details

(a) Task A

Table 22a shows that there is a statistically significant difference between the means of the schools' performance. Nonetheless, the means separation, (Table 22b) indicates that the difference between the means of schools A and B on one hand and A and C on the other is significant.

Table 22a: Analysis of Variance of Skills Exhibited on Task A

Task A	Sum of Squares	Df	Mean Square	F	Sig.
Between	29.678	2	14.839	5.092	0.007

Groups					
Within Groups	515.850	177	2.914		
Total	545.528	179			

Table 22b: Comparisons of Means of the Various Schools on Task A Using Tukey HSD

(I) school	(J) school	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
school A	school B	0.900[*]	0.312	0.012	0.16	1.64
	school C	0.817[*]	0.312	0.026	0.08	1.55
school B	school A	-0.900[*]	0.312	0.012	-1.64	-0.16
	school C	-0.083	0.312	0.961	-0.82	0.65
school C	school A	-0.817[*]	0.312	0.026	-1.55	-0.80
	school B	0.083	0.312	0.961	-0.65	0.82

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

(b) Task B
Table 23a shows that there is a statistically significant difference between the means of the schools' performance. Nonetheless, the means separation, (Table 23b) shows that the difference between the means of schools A and B on one hand and A and C on the other is significant.

Table 23a: Analysis of Variance of Skills Exhibited on Task B

Task B	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	604.744	2	302.372	11.255	0.000
Within Groups	4755.367	177	26.866		
Total	5360.11	179			

Table 23b: Comparisons of Means of the Various Schools on Task B Using Tukey HSD

(I) school	(J) school	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
school A	school B	4.383[*]	0.946	0.000	2.15	6.62
	school C	0.033[*]	0.946	0.005	0.80	5.27
school B	school A	-4.383[*]	0.946	0.000	-6.62	-2.15
	school C	-1.350	0.946	0.329	-3.59	0.89
school C	school A	-0.033[*]	0.946	0.005	-5.27	-0.80
	school B	1.350	0.946	0.329	-0.89	3.59

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

(c) Task C
Table 24a shows that there is a statistically significant difference between the means of the schools' performance. Nevertheless, the means separation, as shown in Table 24b, indicates that the difference between the means of schools A and B is significant.

Table 24a: Analysis of Variance of Skills Exhibited on Task C

Task C	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	29.344	2	14.672	4.143	0.017
Within Groups	626.767	177	3.541		
Total	656.11	179			

Table 24b: Comparisons of Means of the Various Schools on Task C Using Tukey HSD

(I) school	(J) school	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
school A	school B	0.983*	0.344	0.013	0.17	1.80
	school C	0.583	0.344	0.209	-0.23	1.40
school B	school A	-0.983*	0.344	0.013	-1.80	-0.17
	school C	-0.400	0.344	0.476	-1.21	0.41
school C	school A	-0.583	0.344	0.209	-1.40	0.23
	school B	0.400	0.344	0.476	-0.41	1.21
*. The mean difference is significant at the 0.05 level.						

In this study, school type refers to SHS offering General Science (school A), secondary/technical school (school B) and SHS that does not offer General Science nor secondary/technical (school C). The study therefore sought to assess students planning skills in the performance of some tasks (distillation, density and osmosis). Their performances were statistically analyzed based on the school types.

A one-way between-group analysis of variance was conducted to explore the impact of school type on the performance of the three tasks. Students were grouped into three based on the schools they were sampled from. Analysis of performance of school type on task A indicates that there was a statistically significant difference at the $p < 0.05$ level in performance scores for the three groups [F (2, 177) = 5.092, p = 0.007]. Despite reaching statistically significance, the actual difference in mean scores between groups was quite moderate (Cohen, 1988) based on the calculated effect size of 0.054, using eta squared. Post-hoc comparisons using the Tukey HSD test indicated that the mean score for school A [M = 5.93, SD = 1.177] was significantly different from school B [M = 5.03, SD = 1.974]. Also the mean score for school A was significantly different from school C [M = 5.12, SD = 1.860]. However, school B did not differ significantly from school C.

Analysis of performance of school type on task B indicates that there was a statistically significant difference at the $p < 0.05$ level in performance scores for the three groups [F (2, 177) = 11.255, p = 0.000]. The effect size, calculated using eta squared was 0.113. Post-hoc comparisons using the Tukey HSD test indicated that the mean score for school A [M = 12.75, SD = 4.221] was significantly different from school B [M = 8.37, SD = 5.840]. Again the mean score for school A was significantly different from school C [M = 9.72, SD = 5.355]. Nonetheless, school B did not differ significantly from school C.

Analysis of performance of school type on task C indicated that there was a statistically significant difference at the $p < 0.05$ level in performance scores for the three groups [F (2, 177) = 4.143, p = 0.017]. However, the actual difference in mean scores between groups was very small. The effect size, calculated using eta squared was 0.000. Post-hoc comparisons using the Tukey HSD test indicated that the mean score for school A [M = 6.80, SD = 1.592] was significantly different from school B [M =5.82, SD = 2.063]. The mean score for school A was not significantly different from school C [M = 6.22, SD = 1.958]. On the other hand, school B did not differ significantly from school C.

On the whole, students of school A did significantly better than the other schools in all the three tasks because they probably enjoyed enhanced teaching amenities (thus well-established laboratory) and the services of teachers teaching General Science thus trained. This directly supports the findings of Beaumont-Walters and Soyibo, (2001) and Soyibo and Johnson (1998) who observed that students could perform better when they receive better facilities and services of teachers of better quality.

In summary, the findings bring greater insights to the understanding of student performance on planning skills in Integrated Science. On the whole students performed relatively better on the planning skill. Nonetheless there was a comparable performance between males and females. There were significant differences between the performances of the various schools. The students generally, failed to state safety measures, zero balance and state formula for calculating density. With the notable exception of these, the students were able to display the expected skills required for the performance of the various tasks.

Summary of Major Findings

1. It was found out from the study that generally, students were able to outline the various steps required to execute particular tasks.

2. Only a few of students were able to state safety measures for all the three tasks.

3. The task related to materials that students were familiar with was well answered.

4. Most of the students could not zero the weighing balance before using it.

5. It was observed that there was a significant difference in the performance of males and females on task A for school C where females outperformed their male counterparts.

6. There was no significant difference in the performance of male and females on task A for schools A and B.

7. On the whole, there were more similarities than differences between the performances of females and males on of planning skills of the various tasks.

8. It was observed that there was a statistically significant difference between the means of the various schools on task B. The mean score for school A of tasks B was statistically significantly different from schools B and C.

9. There was no statistically significant difference between schools B and C for all the three tasks.

10. There was no statistically significant difference between schools A and C for tasks A and C.

Conclusions

Performance-based assessment is an alternate way of assessing student's planning skill which is related to real life situation. Students displayed their proficiency level in planning the various tasks. The findings negate the statement that a great number of students strategy for planning experiments, are often unsystematic (Kuhn & Phelps, 1982; Hammann, et al., 2008).

The findings on gender support that of Klein et al. (1997); Pine et al., (2006) and Shaw and Nagashima (2009) who claimed that females outperformed their male counterparts in performance assessment tasks. The students however, performed relatively better on the skill of recording values. The current study corroborates the findings of Beaumont-Walters & Soyibo (2001) who observed that students performed relatively better on the skill of recording data. The findings also support the findings of Ssempala (2008) who asserted that there are more similarities than differences between the performances of females and males on performance assessment tasks.

The findings on school type indicated that there were statistically significant differences between the various school types and performance on tasks B. The performances of students of school A were not statistically significant different from schools C on tasks A and C. However, the performance of students of school B did not differ significantly from school C on all the tasks.

Recommendations

Based on the key findings of the study it is recommended that:

Schools should be well equipped with science apparatus for students to interact with materials during science lessons.

Teachers should make the effort to encourage students to practice safety measures when caring out science activities.

Teachers should guide students on how to zero a mechanical balance before using it in measuring the mass of substance for accurate results. This is due to the fact that most of the indicators on the (pointer) balance does not lie on the zero mark when not in use.

Test of practical questions should be related real life situations. This makes science concepts more meaningful to students

Teachers should use materials that students are familiar with within the environment during test of practical lessons for students to really appreciate the nature of science.

Suggestion for Further Studies

This study is limited in scope; the data come from only one of the categories of school classification by the Ministry of Education, in one out of 201 districts. Future studies can expand the scope to include other categories of schools and the various school types.

References

[1] Addai, J. K. (2001). *Evaluating practical skills of Senior Secondary School students with some non-traditional Tasks in mechanics*. Unpublished Master's Thesis: University of Cape Coast, Cape-Coast.

[2] Aikenhead, G. S. (2001). Students' ease in crossing cultural borders into school science. *Science Education*, *85*, 180 –188.

[3] Ainley, M., Hidi, S., & Berndorff, D. (2002). Interest, learning, and the psychological processes that mediate their relationship. *Journal of Educational Psychology, 94*(3), 545 – 561.

[4] Akerson, V., Morrison, J., & McDuffie, A. (2002). *Performance assessment in science methods: a program description*. Retrieved September 12,
   a. 2010, from www.ed.psu.edu/CI/Journals/2002aets/f7_akerson_morrison_m.rtf

[5] Akpan, J. P. (2002). Which comes first: computer simulation of dissection or a traditional laboratory practical method of dissection. *Electronic Journal of Science Education*, *6*(4), 1 – 10.

[6] Alsadaawi, A. (2008). *An investigation of performance-based assessment in science in Saudi primary schools.* Paper presented at the Annual Conference of the International Association for Educational Assessment, Cambridge, 7 – 12 September

[7] Ameyaw-Akumfi, C. (2004). *The promotion of science education in Ghana*. A paper presented at the first national forum on harnessing research, science and technology for sustainable development in Ghana, held at the Accra International Conference Centre.

[8] Anthony-Krueger, C. (2001). *Assessing some process skills of senior secondary school students in elective Biology*. Unpublished Master's Thesis: University of Cape Coast, Cape-Coast.

[9] Ates, S. (2005). The effectiveness of the learning-cycle method on teaching DC circuits to prospective female and male science teachers. *Research in Science and Technological Education*, *23*(2), 213 – 227.

[10] Atkin, J. M., Black, P., & Coffey, J. (2001). *Classroom assessment and the national science education standards*. Washington, DC: National Academies Press.

[11] Baker, E. L., O'Neil Jr., H. F., & Linn, L. (1993). Policy and validity prospects for performance-based assessment. *American Psychologist, 48*(12), 1210 – 1218.

[12] Bandura, A. (1997). Self-efficacy. *Harvard Mental Health Letter, 13*(9), 4-6.

[13] Baser, M., & Durmus, S. (2010). The effectiveness of computer supported versus real laboratory inquiry learning environments on the understanding of direct current electricity among pre-

service   elementary school teachers. *Eurasia Journal of Mathematics, Science   and Technology Education, 6*(1), 47 – 61.

[14] Baxter, G. P., Shavelson, R. J., Goldman, S. R., & Pine, J. (1992). Evaluation   of   procedure-based scoring for hands-on science assessment. *Journal   of Educational Measurement, 29*, 1-17.

[15] Beaumont-Walters, Y., & Soyibo, K. (2001). An analysis of high school   students' performance on five integrated. *Research in Science and   Technological Education, 19*(2), 133 – 145.

[16] Bekiroglu F. O. (2008). Performance based assessments: theory and practice.   *Journal of Turkish Science Education, 5*(1), 113 – 131.

[17] Bell, R. L. (2009). *Teaching the nature of science*. Retrieved February 10, 2011   from www.ngsp.comPortals/0/downloads/SCL220449A_AM_Be   ll.pdf

[18] Biondi, L. (2001). Authentic assessment strategies in fourth grade. *(ERIC   Document   Reproduction Service No. ED460165)*

[19] Bluman, A. G. (2004). *Elementary statistics: a step by step approach*. Boston: Mc Graw Hill.

[20] Brann, A., Gray, T., Piety, J. P., & Silver-Pacuilla, H (2010). *Using   Technology to Support Struggling Students in Science*. Centre for   Implementing Technology in Education. Retrieved September 17,   2010 from www. cited.org

[21] Birmingham, P., & Wilkinson, D., (2003). *Using research instruments a guide   for   researchers* London, Routledge.

[22] Brown, J. D. (2004). Performance assessment: existing literature and   directions   for   research. *Second Language Studies, 22*(2), 91 – 108.

[23] Brown, J. H., & Shavelson, R. (1996). *Assessing hands-on science*: *A   teacher's   guide   to performance assessment*. Thousand Oaks: CA,   Corwin Press.

[24] Buchele, S. (2008). Southwestern professor returns from two years of   teaching   in   Ghana   on fulbright award. Retrieved February 3, 2011 from http://www.southwestern.edu/newsroom/story.php?id=110

[25] Buhagiar, M. A. (2007) Classroom assessment within the alternative   assessment   paradigm: revisiting the territory. *The Curriculum Journal,   18*(1), 39 – 56.

[26] Carin, A. A., & Bass, J. E. (1997). *Methods for teaching science as inquiry*   (8th   ed.).   New Jersey: Merril Prentice Hall.

[27] Century, D. N. (2002). *Alternative and traditional assessments: their   comparative   impact   on students' attitudes and science learning   outcomes: an exploratory study.* Unpublished Doctoral dissertation,   University of Temple, USA.

[28] Chabalengula, V. M., Mumba, F., Hunter, W. F. J., & Wilson, E. (2009). A   model   for assessing students' science process skills during science lab   work. *Problems of Education in the 21st Century, 11,* 28 – 36.

[29] Cheng, M. M. (2008). Identifying strategies to support junior secondary   students   to   engage   in scientific investigation tasks. *Canadian Journal   of Science, Mathematics, and Technology Education*, *8*(2), 99–120.

[30] Chinn, C., & Brewer, W. (1986). An empirical test of a taxonomy of   responses   to   anomalous data in science. *Journal of Research in Science   Teaching, 35*(6), 623 – 654.

[31] Cohen, J. (1988). *Statistical power analysis for the behavioural science*.   Hillsdale, NJ: Erlbaum.

[32] Colvill, M., & Pattie, I. (2002). Science skills – the building blocks for   scientific   literacy. Investigating: *Australian Primary and Junior   Science Journal, 18*(3), 20 – 31.

[33] Corcoran, C. A., Dershimer, E. L., & Tichenor, M. S. (2004). A teacher's   guide   to   alternative assessment taking the first steps. *Clearing House,   77*(5), 213 – 216.

[34] Darling-Hammond, L. & Adamson, F. (2010). *Beyond basic skills: The role*
    a. *of performance assessment in achieving 21st century standards of   learning.* Stanford,   CA: Stanford University, Stanford Center for   Opportunity Policy in Education.

[35] Dooley, K. E., Linder, J. R., & Dooley L. M. (2005). *Advanced methods in   distance education*: *application and practice for educators, trainers   and learners*. London: Idea Group Inc.

[36] Doran, R. L., Chan, F., Tamir, P., & Lenhardts, C. (2002). *Science educators   guide   to laboratory assessment* (2nd ed.). Arlington: NSTA   press.

[37] Dowd, R. (2003). How national competency exams affect International   political   culture. *Journal of Public and International Affairs*, *14*, 1 –   24.

[38] Downing, S. M., & Haladyna, T. M. (2006). *Handbook of text development*. Lawrence Erlbaum Associates.

[39] Enger, S. K., & Yager, R .E. (2001). *Assessing student understanding in science: a standards-based K-12 handbook*. Thousand Oaks, CA: Etkinaorwin Press, INC.

[40] Etkina, E., Karelina, A., & Ruibal-Villasenor, M. (2008). How long does it take? A study of students' acquisition of scientific abilities. *Physical Review Special Topics – Physics Education Research,4*(202108), 1-15

[41] Fairbairn, S. (2007). Facilitating Greater Test Success for English Language Learners. *Practical Assessment, Research & Evaluation, 12*(11), 1 – 7.

[42] Fredua-Kwarteng, Y., & Ahia, F. (2005). Ghana Flunks Math and Science: Analysis (2). Retrieved February 3, 2011 from www.modernghana.comnews/72489/1/ ghana-flunks-math-and-science-analysis-2.html

[43] Garbus, P. (2000). Performance-based assessment in the days of standardized tests, for the Massachusetts Charter School*:* Fellowship Program. Retrieved February 2, 2011 from www.masscharterschools.org/ fellowships/docs/000056/index.html

[44] Gobert, J. D., Pallant, A. R., Krach, N., & Daniels, J. T. M. (2010). Unpacking inquiry skills from content knowledge in geoscience: A research and development study with implications for assessment design Social Sciences and Policy Studies. Retrieved October 27, 2010 from www.wpi.edu/academics/facultydir/jg1.html.

[45] Graves, D. H. (2002). *Testing is not teaching.* Portsmouth, NH: Heinemann.

[46] Grumbine, R. (2010). *R*esearched into using data-collection activities to enrich science courses. *The American Biology Teacher, 72*(6), 369 – 372.

[47] Guy, D., & Wilcox, J. (2000). *Assessing elementary science teaching in a performance setting.* Retrieved September 14, 2010 from www.ed.psu.edu/CI/Journals/2000AETS/18guy_wilcox.rtf

[48] Haigh, M., France, B., & Forret, M. (2005). Is 'doing Science' in New Zealand classrooms an expression of scientific inquiry? *International Journal of Science Education, 27*(2), 215 – 226.

[49] Hammann, M., Phan, T. T. H., Ehmer, M., & Grimm, T. (2008). Assessing pupils' skills in experimentation. *Educational Research, 42*(2), 66 – 72

[50] Hand, B., & Prain, V. (2002). Teachers implementing writing-to-learn strategies in junior secondary science: A case study. *Science Education 86*, 737– 755.

[51] Hand, B., Hohenshell, L., & Prain, V. (2004). Exploring students' responses to conceptual questions when engaged with planned writing experiences: A study with year 10 science students. *Journal of Research in Science Teaching*, *41*(2), 186 – 210.

[52] Hiedi, R. (2008). Performance assessment links in science. Retrieved February 9, 2011 from www.learningreviews.com/Teacher-Tools/Evaluation- Assessment/Performance Assessment-Links-in-Science.html

[53] Herrington, J., & Kervin, L. (2007). Authentic learning supported by technology: Ten suggestions and cases of integration in classrooms. *Educational Media International*, *44*(3), 219 – 236.

[54] Hill, T., & Lewicki, P. (2007). *Statistics methods and applications.* Tulsa: StatSoft.

[55] Hofstein, A., Shore, R., & Kipnis, M. (2004). Providing high school chemistry students with opportunities to develop learning skills in an inquiry-type laboratory: a Case Study. *International Journal of Science Education, 26*(1), 47 – 62.

[56] Huff, J. (1998). Assessment in science: an examination of multiple-choice and performance science assessment for a second grade classroom. *(ERIC Document Reproduction Service No. ED423286)*

[57] Jacobs, R. M (2010). Educational research: Causal-comparative studies. EDU 8603. Retrieved August 16, 2011 from
  a. www83.homepage.villanova.edu/richard.jacobs/.../lessons/causal.ppt

[58] Janisch, C., Liu, X., & Akrofi, A. (2007). Implementing alternative assessment: opportunities and obstacles. *The Educational Forum,* 71, 221-230

[59] Johnson, E. A. (2001). *Assessing laboratory skills of Physics students in selected senior secondary school topics in mechanics*. Unpublished Master's Thesis: University of Cape Coast, Cape-Coast.

[60] Jovanovic, J., & King, S. S. (1998). Boy and girls in the performance-based science classroom: who's doing the performing? *American Educational Research Journal, 35*(5), 477 – 496.

[61] Jovanovic, J., Solano-Flores, G., & Shavelson, R. J. (1994). Performance- based assessments: Will gender differences in science achievement be eliminated? *Education and Urban Society, 26* (4), 352 – 366.

[62] Kabba, C. (2008). Performance-based assessment (Idea Bank: Tips and Techniques for Creative Teaching). *The Science Teacher,75*(8), 68 – 71

[63] Kajornboon, A. B (2005). Using interviews as research instruments. Retrieved August 16, 2011 from www.culi.chula.ac.th/e-Journal/bod/Annabel.pdf

[64] Karelina, A., & Etkina, E. (2007). Acting like a physicist: Student approach study to experimental design. *Physical Review Special Topics Physics Education Research.* 3, 020106, Retrieved December 10, 2010 from URL:http://link.aps.org/doi/10.1103/PhysRevSTPER.3.020106

[65] Kathy, S. (2010). Alternative/performance-based assessment. Retrieved 26 November, 2010 from www.emtech.net/Alternative_Assessment.html

[66] Keli, C., Haney, J., & Zoffel, J. (2009). Improvement in students achievement and science process skills using environmental health science problem- based learning curricula. *Electronic Journal of Science Education* 13(1), 1 – 18.

[67] Kemi, J., & Adsit, J. (2008). *Goals, guidelines and standards for scientific investigations.* North American council for online learning. Retrieved February 4, 2011 from http://www.inacol.org.

[68] Ketelhut, D. J. (2007). The impact of student self-efficacy on scientific inquiry skills: An exploratory investigation in River City, a multi-user virtual environment. *The Journal of Science Education and Technology,* 16(1), 99 – 111.

[69] Kilfeather, P., O'Leary, M., & Varley, J. (2006). Adapting science performance tasks developed in different countries for use in Irish primary schools. *Irish Educational Studies, 25*(1), 3 – 33

[70] Klassen, S. (2006). Contextual assessment in science education: background, issues, and policy. (Electronic Version). *Wiley InterScience.* Retrivied April 8, 2011, from www.interscience.wiley.com

[71] 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. *Educational Evaluation and Policy Analysis, 19*(2), 83 – 97.

[72] Kohlhaas, K., Lin, H. & Chu, K. (2010). Disaggregated Outcomes of Gender, Ethnicity, and Povertyon Fifth Grade Science Performance. *Research in Middle Level Education Online, 33*(7). 1-12

[73] Kuhn, D., & Phelps, E. (1982). The development of problem-solving strategies. In: H. Reece (Ed.), *Advances in Child Development and Behaviour.* New York: Academic Press

[74] Lee-Ann, F. (2008). A teacher's outlook on testing is changed after students are assessed through hands-on Tasks. *Journal of Science and Children,* 39(9), 32 – 35.

[75] Leonard, W. H., Speziale, B. J., & Penick, J. E. (2001). Performance assessment of a standards-based high school biology curriculum. *The American Biology Teacher, 63*(5), 310 – 316.

[76] Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance- based assessment: expectations and validation criteria. *Educational Research, 20*(8), 15 – 21

[77] Liu, J. (2000). *The effect of performance-based assessment on eighth grade students Mathematics achievement.* Unpublished Ph.D dissertation. University of Missouri; Columbia.

[78] Lynch, S. J. (2000). *Equity and science education reform.* Mahwah: Lawrence Erlbaum Associates, Inc.

[79] Mason, R. D., Lind, D. A., & Marchal, W. G. (1996). *Statistical techniques in business and economics.* Boston, Irwin McGraw-Hill.

[80] Myers, B. E & Dyer, J. E (2004). Assessing agriculture teachers' capacity for teaching science integrated process skills. *Journal of Southern Agricultural Education Research, 54*(1), 74-84

[81] Mayer-Smitha, J., Pedrettib, E., & Woodrowa, J. (2000). Closing of the gender gap in technology enriched science education: a case study. *Computers and Education, 35*, 51 – 63.

[82] Ministry of Education Science and Sports (MOESS). (2007). *Teaching syllabus for Integrated Science (SHS).* Accra: Curriculum Research and Development Division.

---

[83] Morrison, J., McDuffie, A., & Akerson, V. (2003). Preservice teachers' development and implementation of science performance assessment tasks. *ERIC Document Reproduction Service No. ED 478065.*

[84] Mkpughe, M. L. (1998). *The interaction of gender, location, and socio- economic status on students' academic performance in home economics at the junior secondary school level.* Unpublished M.Ed dissertation, Delta State University; Nigeria

[85] National Academy of Science. (1996). *National science education standards.* Washington, DC: National Academy Press.

[86] North Central Regional Educational Laboratory (2005).Critical Issue: Using Technology to Improve Student Achievement. Retrieved August 16, 2011 from http://www.ncrel.org/sdrs/areas/issues/methods/technlgy/te800.htm

[87] Office of Technology Assessment (OTA). (1992). *Testing in American schools: Asking the right questions.* Washington, DC: OTA.

[88] Okoye, N. S. (2009). The effect of gender, socio-economic status and school location on students' performance in Nigerian integrated science. *Education, 129*(3), 561 -568

[89] Oloruntegbe, K. O. (2010). Approaches to the assessment of science process skills: a reconceptualist view and option. *Journal of College Teaching & Learning, 7*(6), 11 – 18

[90] Oloruntegbe, K. O. (2000). *Effect of teachers' sensitization on students' acquisition of science process skills and attitudes,* Unpublished Ph.D Thesis of Faculty of Education, University of Benin

[91] 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.

[92] Onyango, M. O. (2008). Assessment and teaching science. *The International Journal of Learning, 15*(4), 256 – 259

[93] Ossei-Anto, T. A. (1996). *Assessing laboratory skills of students in selected High School Physics topics in optics.* Unpublished Ph.D dissertation. State University of New York; Buffalo

[94] Padilla, M. J. (1990). *The science process skills. Research Matters - to the Science secondary schools* (5th ed.). Upper Saddle River, N.J: Merrill Prentice Hall.

[95] Parker, A. V., & Gerber, B. L. (2002). Performance based assessment, science festival exhibition presentations, and elementary science achievement. *Journal of Elementary Science Education, 14*(1), 59 – 67.

[96] Perlman, C. C. (2003). Performance Assessment: Designing Appropriate Performance Tasks and Scoring Rubrics. *ERIC Document Reproduction Service No. ED480070.*

[97] Pine, J., Aschbacher, P., Roth, E., Jones, M., McPhee, C., Martin, C., Phelps, S., Kyle, T., & Foley, B. (2006). Fifth graders' science inquiry abilities: A comparative study of students in hands-on and textbook curricula. *Journal of Research in Science Teaching, 43*(5), 467 – 484

[98] Pico II, R. A. (1999). *A case study of three Puerto Rico-statewide systemic initiative intermediate science teachers' views on and use of performance assessment.* Unpublished doctoral dissertation, The Pennsylvania State University, Malvern

[99] Popham, W. J. (2000). *Modern educational measurement* (3rd ed.). Needham Heights, MA: Allyn & Bacon.

[100] Quellmalz, E. S., Hinojosa, T., & Padilla, C. (1999). Performance assessment links in science (Electronic version). *Practical Assessment, Research and Evaluation, 6*(10). Retrieved April 5, 2010 from http://PAREonline.net/getvn.asp?v=6&n=10.

[101] Rambuda, A. M., & Fraser, W. J. (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.

[102] Rezba, R. J., Sprangue, C., Fiel, R. L., & Funk, H. J. (1995). *Learning and assessing science process skills* (3rd ed.). Dubuque, Iowa: Kendall/Hunt Publishing Company.

[103] Roth, W., Eijck, V. M., Hsu, P., Marshall, A., & Mazumber, A. (2009). What high school students learn during internships in biology laboratories: *The American Biology Teacher, 71*(8), 479 – 493.

[104] Sarantakos, S. (1988). *Social research* (2nd ed.). Charles Sturt University, Australia: Palgrave

[105]   Scott, S. J. (2002). *Using multiple methods of assessment to examine student   achievement   in Music*.   Retrieved   February   1,   2011   from http://www.brandonu.ca/Music/People/FacultyPages/scotts/using_mult iple_methods_of_assess.htm

[106]   Seiler, G. (2001). Reversing the 'standard' direction: Science emerging from   the   lives   of African American students. *Journal of Research in Science Teaching, 38*(9), 1000 – 1014.

[107]   Seshie, A. H. K. (2001). A*ssessing laboratory skills of students in selected   senior secondary school elective chemistry topics in titrimetric   analysis*. Unpublished Master's Thesis: University of Cape   Coast, Cape-Coast.

[108]   Shavelson, J. R., Baxter, P. G., & Pine, J. (1991). Educational research.   *Applied Measurement in Education, 4*(4), 347 – 362.

[109]   Shavelson, R. J., Baxter, G. P., & Pine, J. (1999), Performance Assessment in   Science. *Applied Measurement in Education*, *4*(4), 347 – 362.

[110]   Shaw, J. M., &. Nagashima, S. O. (2009). The achievement of student   subgroups   on   science performance assessments in inquiry-based   classrooms. *Electronic   Journal   of   Science Education,13*(2), 6 - 29.

[111]   Shymansky, J. A., Chidsey, J. L., Henriques, L. E., Yore, L. D. S., Wolfe, E.   W., & Jorgensen, M. (1997). Performance assessment in science as a tool to enhance the picture of student learning. *School Science and   Mathematics, 97*(4), 172 - 183.

[112]   Slater, T., F. (1997). The effectiveness of portfolio assessments in science.   *Journal of College Science Teaching, 26*(5), 1-10

[113]   Soyibo, K., & Johnson, R. (1998). Science knowledge, science attitudes and   self-esteem: a comparison of ROSE and non-ROSE grades 7 and 8   students, *Caribbean   Journal   of Education*, *19*, 163 - 178.

[114]   Spektor-Levy, O., Eylon, B., & Scherz, Z. (2009). Teaching scientific   communication   skills   in science studies: does it make a difference? *International Journal of science and Mathematics Education*, *7*, 875–   903

[115]   Ssempala, F. (2008). *Gender differences in performance of chemistry practical   skills   among senior six students in Kampala district*. Unpublished   doctoral dissertation. Dissertation.com, Boca Raton, Florida, USA

[116]   Sutherland, L. M., Shin, N., & Krajcik, J. S. (2010). *Exploring the Relationship   between   21st Century Competencies and Core Science   Content*. Retrieved October 27, 2010 from ww7.nationalacademies.org/.../Sutherland_Shin_Krajcik_Core_Scien   ce Content_Paper 1- 22.

[117]   Tachie, T. Y. (2001). *Assessing observational skills of junior secondary   school students   in   the Ashiedu Keteke Sub Metro of the Accra   Metropolis*. Unpublished Master's Thesis: University of Cape Coast,   Cape-Coast

[118]   Tai, R. H., Liu, C. Q., Maltese, A. V., & Fan, X. (2006). Planning early for   careers in science. *Science*, *312*, 1143–1144.

[119]   Tan, S. C., Yeo, A. C. J., & Lim, W. Y. (2005). Changing epistemology of   science   learning through inquiry with computer-supported   collaborative   learning. *Journal   of   Computers   in Mathematics and   Science Teaching*, *24*(4), 377–386.

[120]   Teachnology (2011) The Online Teacher Resource: Constructivism learning   theory. Retrieved August 16 2011, from   http://www.teachnology.com/search/search.cgi?zoom_query=author&s a.x=6 &sa.y=24

[121]   Trowbridge, L. W., Bybee, R. W., & Powell, J. C. (2000). *Teaching secondary   school   science: strategies for developing scientific literacy* (7th ed.). New Jersey: Merril Prentice Hall

[122]   Tung, R., & Stazesky, P. (2010). *Including performance assessments in   accountability systems: A review of scale up efforts*. Center for Collaborative Education and the Nellie Mae Education.

[123]   Tüysüz, C., Karakuyu, Y., & Tatar, E. (2010). Opinions of students' parents   about performance Task in science and technology class. *Necatibey   Faculty of Education Electronic Journal of Science and Mathematics   Education, 4*(1), 108 – 112.

[124]   Ukwungwu, J. O. (2002), Gender difference study of performance in   Integrated   Science.   A summary of studies conducted in Nigeria.   *Journal of Science Teachers Association of Nigeria, 37*(2), 55 – 59.

[125]   Walker, K. (2010). *What Should Parents Know About Performance Based Assessment.*   Retrieved January 12, 2011 from   http://www.projectappleseed.org/assessment.html

[126]   West African Examination Council (WAEC) (2004). *Chief Examiners'   Report   for   Senior Secondary Certificate Examination, Integrated   Science.* Accra: WAEC.

[127]   West African Examination Council (WAEC) (2005). *Chief Examiners'   Report   for   Senior Secondary Certificate Examination, Integrated   Science.* Accra: WAEC.

[128]   West African Examination Council (WAEC) (2006). *Chief Examiners'   Report   for   Senior Secondary Certificate Examination, Integrated   Science.* Accra: WAEC.

[129]   West African Examination Council (WAEC) (2007). *Chief Examiners'   Report   for   Senior Secondary Certificate Examination, Integrated   Science*. Accra: WAEC.

[130]   Williams, D. D., Howell, S. L., & Hricko, M. (Eds.). (2006). *Online   assessment   and measurement: Foundations and challenges*. Hershey: Information Science Publishing.

[131]   Winking, D., & Bond, L. (1995). *What you and your school should know   about   alternative assessment*. Oak Brook, IL: North Central Regional Educational Laboratory.

[132]   Wren D. G. (2009). *Performance assessment: a key component of a balanced   assessment system.* Report from the Department of Research,   Evaluation, and Assessment.

[133]   Yujing, N. (1997). Performance-Based Assessment: Problems and Design Strategies.   *Education Journal, 25*(2), 137-157