Effects of Two-Tier Multiple Choice Diagnostic Assessment Items on Students' Learning Outcome in Basic Science Technology (BST)

S O Adodo Ph.D

Science and Technical Education Department, Adekunle Ajasin University,
Akungba Akoko. P.M B 001, Ondo State, Nigeria.
E-mail: so_adodo@yahoo.com, Tel.:+2348035064033

Abstract

This paper addresses the effect of The TTMC diagnostic assessment Model in basic science technology and to see if it can make an important contribution to the student learning outcome in basic science and technology among Nigerian students at the junior secondary school (JSS) level. The paper addresses the importance of the diagnostic assessment in basic science and technology and a review of development in particular the TTMC diagnostic instrument and sees if it can make any contribution to students' learning outcome among Nigerian students at the Junior Secondary School (JSS) level. The study is a quasi-experimenter research designed to determine the effect of TTMC diagnostic instruction on students' performance in BST. It is quasi experiment study in which the effect of the treatment on interest, attitude and achievement was measure as students' learning outcome. The sample of the study consisted of 108 JSS basic science students selected randomly from 3 junior secondary schools in Ondo state Nigeria. Among these students were 54 male and 54 females selected through stratified random sampling technique. The instrument for dada collection in the study were BST Two-tiers multiple choice TTMC achievement test of 20 items adopted from Tan and Trannguest 1990; Tin 2004 Wang 2004; and from Chin and Ho 2002; The test was administered to the subject as pre- test; before the treatment and as post-test immediately after the treatment and the retention test after 4 weeks of the post-test except for the control group which was not expose to any treatment. The scores were analysed using mean; standard deviation; analysis of variance ANOVA; ANCOVA and the multiple classification analysis.

1. Introduction

The ultimate aims of science education in every society are to provide people with knowledge of science needed for the fulfilment of the socio-economic and cultural needs of the society as well as ensure the means of physical survival. The federal Government of Nigeria seem to live effectively in view and having realised these by stating in the National Policy on education, that one of the aim of education should be equipping student's to live effectively in our Modern age of science and technology (NPE 2004).

In view of these aims, the policy emphasised the need for effective teaching and learning of science in the primary schools as basic science and technology up to the first three years in the secondary school Junior Secondary School (JSS) before they are exposed to the compartmentalised science (Physics, chemistry and Biology) at the senior secondary and the post-secondary institutions.

The nature and the extent of students understanding of scientific concepts and phenomena are key component of any sciences curriculum. In order to gauge the effectiveness of classroom instruction to facilitate students understanding of scientific concepts, assessment test have to be readily available for use by classroom teachers. A large body of outcome shows that majority of teachers do not effectively diagnose students learning problems, especially at an early stage of students learning process (Taber 2011). How teachers can address students learning needs by incorporating in their instructional repertories specially designed assessment procedures should be an integral part of their teaching activities (Gallagher and Parker 2001).

In most of the research reports about curriculum, reform and concerns about assessment, examine and testing are usually presented as refinement to existing technical testing procedure. Nevertheless, there are notable changes from the norm of testing procedures. In order for science teachers pedagogy to be more effective and productive, diagnostic formative assessment approach are needed as the current assessment procedure to an in- class formative assessment used by science teachers in the standard test which are largely paper-pencil collection of individual items with single correct answers MCQs presented without a surrounding context..To this end, a change that will provide opportunity for

teachers to develop clinical judgement about students understanding of significant ideas and processes and encourage student interactive discussion rather than measuring education progress is required.

2. Multiple-Choice questions(MCQs) Tests

Multiple-choice tests have been found an effective way of identifying the misconceptions of the students by researchers. Bar (as cited in Al-Rubayea, 1996) stated that multiple-choice tests are more effective than oral or written open-ended essays in detecting students' misconceptions. Multiple-choice tests have many advantages. They can be scored immediately and objectively. Teacher can administer them easily and they are applicable to large number of students (Al-Rubayea, 1996). Moreover, Çataloglu, 2002 expressed that multiple-choice tests are better liked by the students than other measures and can give diagnostic information. Scott (as cited in Marx, 1988) expressed nine appropriate reasons for using of multiple-choice tests: (1) They provide greater variety of questions. (2) They can be qualitative questions regarding physics principles. (3) Choosing between alternatives and having a general understanding are much more like real life. (4) Options act like hints. (5) The teachers can ask subtle points with them. (6) Multiple-choice items are next best thing to essay type questions. (7) The teachers can ask for a quick numerical calculation and make them worth a point. (8) More material can be covered. (9) They are good for review. There are also some criticisms to the multiple-choice tests. According to Rollnick and Mahooana (1999) the disadvantage of multiple-choice tests is that questions do not provide deep enough inside into the students' ideas on the topic and students very often give correct answers for wrong reasons. According to Çataloglu, 200, multiple-choice tests direct the students' attention on information in isolation by testing one element at a time. Therefore, the larger context and structure of relationships between and among the elements get lost. According to Bork (as cited in Marx, 1988) multiple-choice tests should never be used. He expressed five reasons to support his assertion. Firstly, multiple choice items encourage guessing. Secondly, the items are not from real life situations. Thirdly, that they are not friendly for students. Because, students see them in somewhat a derogatory fashion, connected with the fact that guessing is involved. Sandin (as cited in Marx, 1988) added two more reasons for why multiple-choice tests are not effective: First, students may have extracted the right answer by a fortuitous combination of errors. Second, multiple-choice tests heavily depend on reading comprehension skills. As it is seen, multiple-choice tests are easily applicable and their results can be analyzed quickly and easily but. the problem is their effectiveness. To overcome this problem, (Al-Rubayea, 1996) recommended that students should justify their answers. As a result, researchers extended the multiple-choice tests into several tiers, two or three tiers.

Multiple-choice items(MCQs) is said to have evaluated only content knowledge without considering the reasoning behind students' choices of responses (Duncan and Johnstone, 1973).According to ------, the multiple-choice type have certain drawbacks because they are often compiled by teachers with a correct answer in mind , together with a number of distracters' designed by teachers to give particular types of error. In most of MCQs, the origin of the distracters is not specified and the items do not investigate conceptual understanding . Hence distracters based on students' answers to open-ended questions that will provide better indicators of students' conceptual understanding of specific content area in science is needed. As an improvement of this methodology, Tamir (1971) proposed the use of multiple-choice test items that included responses that also required students to justify their choice of option by giving a reason. Tamir (1989) found the use of justifications when answering multiple-choice test items to be a sensitive and effective way of assessing meaningful learning among students and addresses, to some extent, the limitations of traditional multiple-choice test items. The positive outcomes of findings related to students' justifications to test items led to the development of two-tier multiple-choice diagnostic tests specifically for the purpose of identifying students' alternative conceptions in limited and clearly defined content areas. Treagust (1995) describe a approach using two tiers of multiple choice item with distracters to diagnose students' conceptual understanding of specified content areas in science. The first tier involves a content response and the second tiers s reasoning response. Treagust (1995) has provided useful guidelines for the development of these instruments specifically for the purpose of identifying students' alternative conceptions about various concepts. The first tier as a content question, while the second tier elicits a reasoning response. The need to select a justification in these multiple-choice items affords a sensitive and effective way of assessing meaningful learning among students and also serves as an effective diagnostic tool (Tamir, 1989)

3. Development And Use Of two-tier multiple choice (TTMC) Test Items

A two-tier test is a two level multiple choice question that diagnoses students' alternative conceptions in science (Tsai &

Chou, 2002). The first tier is in the multiple-choice format as king the content knowledge of the students with two, three or four choices. The second tier is again in the multiple-choice format asking the reason for the response given in the first tier Treagust (as cited in Odom & Barrow, 1995). Two-Tier Multiple Choice(TTMC) questions are a more sophisticated form of multiple choice questions. The first tier resembles a traditional MCQ, which usually pertains to a knowledge statement. The second tier resembles the format of a traditional multiple choice question but aim to promote higher thinking and reasoning skills. The questions also aim to identify misconceptions held by students as many of the distracters are based on such misconceptions. Two-tier tests include, in addition to selecting correct answer among the distracters, multiple reasons or justifications from which the students choose their reason for their response is required in the second tiers. Treagust (as cited in Odom & Barrow, 1995) described the item format of the two-tier multiple choice tests as the first tier consisting of a content question with two, three, or four choices. The second tier consists of four possible reasons for the first part with three of them alternative reasons and one desired reason. The second tier can also include a blank that students can write a reason for the first tier when they can not see their reasons among the alternatives of the second tier (Griffard & Wandersee, 2001). Advantages of the two-tier tests

According to Zeilik (n.d.) teachers can use these diagnostic tests for formative and summative assessments over semesters. If teachers use them as a formative test, they will understand their students' cognitive states, preconceptions and misconceptions prior to instruction. Therefore, they can take some precautions for misconceptions which can possibly obstruct the lesson. For example, they can tutor the students in their weak areas individually or assign the students into heterogeneous cooperative learning teams. Or a homogeneous group to be treated separately and specially through re-teaching If teachers use the diag nostic tests for summative assessment, they will see impact of their instruction method positive or negative, which can serve feedback for later on instructions. However, it is important to say that results of the diagnostic tests cannot be used for grouping students. Because, the main purpose of the tests is to diagnose not to assess achievement of the students.

A wide range of specially created two-tier multiple choice instruments have been developed and used to determine students understanding of the concepts in several science disciplines. Examples of instruments developed and used to investigate topics in biology, in chemistry and in physics since 1980's include: Two –tiers diagnostic instrument in Photosynthesis and respiration developed by Haslam and Treagust (1987), in Photosynthesis by Griffard and Wandersee (2001), in Breathing and respiration by Mann and Treagust (1998), in Diffusion and osmosis by Odom and Barrow (1995), in Internal transport in plants and human by Wang (2004) circulatory system Flowering plant growth and development by Lin (2004), Covalent bonding and structure by Treagust and Garnett (1989), Chemical bonding by Tan and Treagust (1999), Qualitative analysis by Tan, Treagust, Goh and Chia (2002), Chemical equilibrium by Tyson, Treagust and Bucat (1999), Multiple representation in chemical reactions by Chandrasegaran,Treagust &.Mocerino (2005), Ionisation energies of elements byTan, Taber, Goh and Chia (2005), Acids and bases by Chiu (2001,2002), Electrical circuits by Millar and Hames (2001) and States of matter by Chiu, Chiu and Ho (2002) Two-Tier (TTMC) have proven to be very successful in educational settings (Williams, 2006) and research has shown that not only do they (1) help to test student understanding, but they also (2) aim to test student higher level of cognitive thinking and (3) help to identify misconceptions students may have (Treagust 2006).

4. Different vivid examples of such two-tier tes (TTMC) as in (Treagust,1995)

4.1 Example 1

Question: The compound formed between magnesium and oxygen can be used as a heat-resistant material to line the walls of furnaces.

    I. True   II. False

Reason

A. The lattice of magnesium oxide resembles that of silicon.
B. The covalent bonds between magnesium and oxygen atoms are strong.
C. The intermolecular forces between the magnesium oxide molecules are weak.
D. There are strong ionic forces between magnesium and oxide ions in the lattice.

The process of construction of two-tier multiple-choice items that test higher level abilities than recall is considered long and difficult. Examples of instruments used to investigate topics in biology, in chemistry and in physics are summarised in Table 1. These two-tier multiple-choice tests are more readily administered and scored than the other

methods of ascertaining students understanding, and thus are particularly useful for classroom teachers enabling them to use the findings of research to inform their teaching. Two-tier items has been used by the National Science Council in Taiwan as the central part of their national assessment project and the American Chemical Society as recommended examples for conceptual questions.

4.2   Example 2 Respiration Questions: by (Lin 2004)

Question: Respiration occurs in both animal cells and plant cells
    A. *True   B. False
    The reason for my answer is because :
    1.   Green plant cells respire when there is no light energy and carry out the process of photosynthesis when there is light energy.
    2.   Respiration only occurs in animal cells and not in plant cells. Plant cells only photosynthesis.
    3.   *Both plant and animal cells respire all the time. Plant cell respiration takes place during the day while photosynthesis is occurring
    4.   Respiration only takes place during the night time hours in plant cells.

4.3   Another question involved in the exploratory study is taken from Boo, (2005)

Question. Study the table carefully

Group	Living things
X	Python, guppy, seahorse
Y	Horse, cat, squirrel

These living things are grouped according to
1)   Where the live
2)   Their outer covering
3)   How they move about
4)   The type of food they eat
Explain your reasoning.

_____
_____
_____
Question 2
Which one of the following animals is unlike the others in the terms of how it moves?
1)   Frogs
2)   Kangaroo
3)   Rabbits
4)   Tiger
Explain the reasoning   (   )

_____
_____
_____

5.   The Use OF (TTMC) Test Items

Tamir (1989) found the use of justifications when answering multiple-choice test items to be a sensitive and effective way of assessing meaningful learning among students and addresses, to some extent, the limitations of traditional multiple-choice test items. As a result, he proposed the use of multiple-choice test items that included responses with known student alternative conceptions, and that also required students to justify their choice of option by giving a reason   The positive outcomes of findings related to students justifications to test items led to the development of two-tier multiple-

choice test items specifically for the purpose of identifying students alternative conceptions is limited and clearly defined content areas. These are short paper and pencil tests that are convenient to administer and not time consuming to mark. The means whereby two-tier items have been designed has being well documented by Treagust and other researchers who have implemented the approach. In brief, there are three major aspects to developing these items: (a) the content is defined by the identification of propositional content knowledge statements of the topic to be taught and the development of a concept map that accommodates the propositional statements; (b) information about students conceptions is obtained from the extent research literature, where available, and where not available by having students provide free response explanations to their answers and conducting unstructured interviews with students who have previously been taught the content/concepts; and (c) the development of the two-tier multiple-choice diagnostic items. Hence the need to select a justification in the multiple-choice items afftords a sensitive and effective way of assessing meaningful learning among students and also serves as an effective diagnostic tool (Tamir 1989).

Problem. The nature and the extent of students understanding of scientific concepts and phenomena are key component of any sciences curriculum. In order to gauge the effectiveness of classroom instruction to facilitate students understanding of scientific concepts, assessment test have to be readily available for use by classroom teachers. A large body of outcome shows that majority of teachers do not effectively diagnose students learning problems, especially at an early stage of students learning process up till when they are tested and assessment is done using the single tier of multiple-choiceMCQs item which usually consists of a content question having usually two to four choices is good but not sufficient to diagnose learners conceptual problem . The construction of two-tier multiple-choice(TTMC) items that test higher level abilities than recall is considered necessary.

Purpose: The paper addresses the importance of the diagnostic assessment in basic science and technology and a review of development in particular the TTMC diagnostic instrument and sees if it can make any contribution to students' learning outcome among Nigerian students at the Junior Secondary School (JSS) level.

Methodology: The study is a quasi-experimenter research designed to determine the effect of TTMC diagnostic instruction on students' performance in BST. It is quasi experiment study in which the effect of the treatment on interest, attitude and achievement was measure as students' learning outcome. The sample of the study consisted of 108 JSS basic science students selected randomly from 3 junior secondary schools in Ondo state Nigeria. Among these students were 54 male and 54 females selected through stratified random sampling technique.

Instrument: The instrument used for the study were BST Two-tiers multiple choiceTTMC achievement test of 20 item adopted from Tan and Trannguest 1990; Tin 2004 Wang 2004; and from Chin and Ho 2002;   The test was administered to the subject as pre- test; before the treatment and as posttest immediately after the treatment and the retention test after 4 weeks of the posttest except for the control group which was not expose to any treatment. The scores were analysed using mean; standard deviation; analysis of variance ANOVA; ANCOVA and the multiple classification analysis.

Research Questions.
1. What is the general performance of students in Basic Science Technology before and after treatment using two-tiers diagnostic model ?
2. What is the level of interest of male and female subjects before and after treatment among the groups

Research hypothesis.

Ho_1 There is no significant differences in the achievement mean scores of the experimental groups and the control group before and after treatment.

6. Results

Question 1: What is the general performance level of students in Basic Science and Technology before and after treatment?

Table 1: Mean scores of the general performance level of students in Basic Science and Technology before and after treatment?

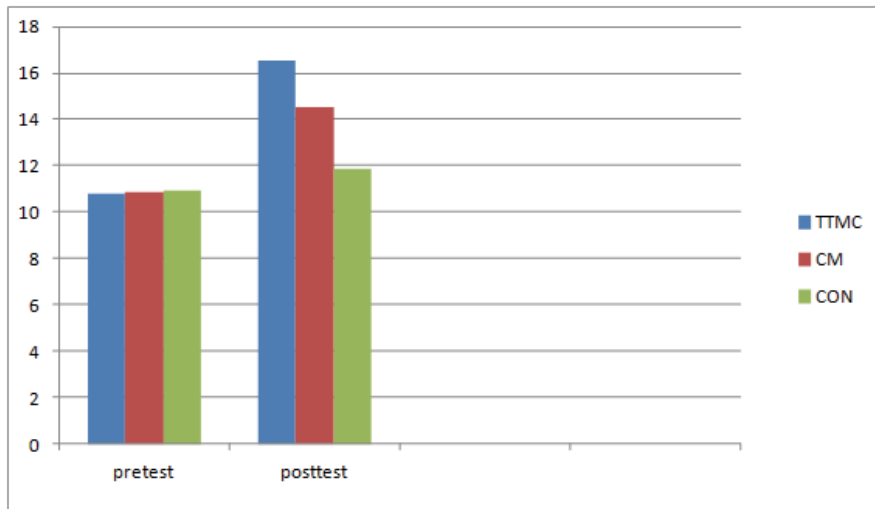| Groups	Pre-test		Post-test Achievement	
	Mean	S.D	Mean	S.D
TTMC	10.81	1.34	16.56	1.64
CM	10.83	1.36	14.50	1.85
Control	10.92	1.37	11.39	2.27

Figure A. Composite graph of Students' performance in Basic Science Technology before and after treatment

The table 2 and the figure A above reveal the performance of the group in the pre-test , post-test and achievement mean scores of the students. The TTMC group had a mean scores of 10.81 , 16.56 and 17.03 in the pre-test , post-test and achievement respectively. While the CM had a mean score of 10.83, 14.50 and 14.26 in the pre-test and post-test and retention respectively. The control group had the mean scores of 10.92 ,11.43 and 11.28 in the pre-test and post-test and retention respectively. The above analysis showed that the three groups use for the study were homogeneous at the beginning of the study. However a closer look at the table showed that there was an increase in the performance of the students among the group after treatment. This fact is supported by the appreciable increase in the achievement mean score of the group., with the TTMC having the highest achievement and retention post-test mean scores.
Question 2: What is the level of interest of male and female subjects before and after treating among the groups.

Table 2: Mean score of interest of male and female subjects before and after treating among

| | MALE | | | | FEMALE | | | |
| | Before | | After | | Before | | After | |
Groups	Mean	S.D	Mean	S.D	Mean	S.D	Mean	S.D
TTMC	23.22	2.55	24.28	2.83	23.61	2.06	24.44	2.27
CM	23.72	2.05	24.19	2.58	23.72	2.33	24.11	2.56
Control	23.00	2.29	23.08	2.28	23.54	2.68	23.46	2.01

The table above showed the interest of the male and female subjects as represented by their mean scores in the Science interest and attitudinal test can be readily seen. For the TTMC group, the boys had the mean score of 23.22 before the treatment and 24.28 (4%) after the treatment. Likewise the girls in the TTMC group had the mean scores of 23.61 and 24.44 (about 3.5%) increase after the treatment. For the CM group, the boys had the mean scores of 23.72 and 24.19(about 2%) increase after treatment. The girls had the mean scores of 23.72 and 24.11(about 1.6%) increase after treatment. However, the boys i the control group had the mean scores of 23.00 and 23.08 in the pre-test and post-test scores respectively (just an increase of .4%) while the girls in the group had mean scores of 23.84 and 23.46 in the pre-test and post-test scores respectively.(about 1.6% decrease)
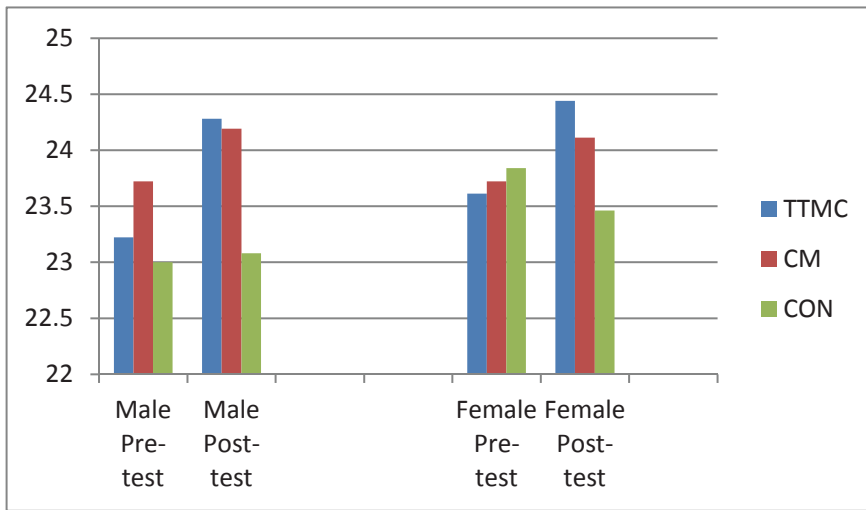
Figure B. Composite graph of Male and Female interest mean scores before and after treatment.

Ho1   There is no significant differences in the achievement mean scores and retention mean scores of the experimental groups and the control group after treatment.

Table 3: Summary of ANCOVA on the achievement and retention Mean Scores of subjects

Source of variation	SS	d.f	MS	F	P
Main Effects	962.532	3	320.844	109.423	.000
Test	.782	1	.782	.267	.606**
Group	961.750	2	480,875	164.001	.000
2-way interaction	8.676	2	4.338	1.479	.230
Test x group	8.676	2	4.338	1.479	.230
Explained	971.208	5	195.242	66.246	.000
Residual	615.730	210	2. 432		
Total	1586.985	215	7.381		

Ho2 There is no significant differences in the retention mean scores of the experimental groups and the control group before and after treatment.

Table 4: To test this hypothesis, a two-way analysis of variance was computed. ANCOVA was also computed for the differences that might exist between the experimental and control groups

Source of variation	SS	d.f	MS	F	P
Covariates	7.285	1	7.285	1.202	.272
Pretest Scores	7.285	1	7.285	1.202	.272
Main Effects	958.482	3	319.494	101.496	.000
Test	.684	1	.684	.259	.526
Group	957.796	2	478,895	161.980	.000
2-way interaction	7.854	2	3.027	1.216	.340
Test x group	7.854	2	3.027	1.216	.340
Explained	966.336	5	193.267	63.452	.000
Residual	620.649	210	2. 956		
Total	1586.985	215	7.381		

Tables 3 and 4 above revealed a significant main effect for the groups (F = 164.001,p <0.000 for ANOVA and F = 161.980, p < 0.000 for ANCOVA but not for tests (F= 0.267,p < 0.606 for ANOVA and (F =0.259,p< 0.526 for ANCOVA. This by implication means that the subjects retained what they learnt. However, to provide some indications of the performance of the group, a multiple Classification Analysis was computed. The result are as presented in table 5.

Table 5: Multiple Classification Analysis (MCA) of achievement and retention mean scores of subjects in the different groups

Grand Mean = 14.236

Variable+ Catergory	N	Unadjusted Deviation	Eta	Adjusted for Independent + Covariate deviation	Beta
Tests					
Achievement	108	.06		.06	
Retention	108	-.06		-.06	
			.02		.020
Groups					
TTMC	36	2.54		2.54	
CM	36	0.08		0.08	
Control	36	-2.63		-2.63	
			.78		.780
Multiple R²					.607
Multiple R					.779

The MCA table above with the grand mean of 14.236 showed that achievement has the higher adjusted mean (14.296). although not significantly different while retention has the lowest adjusted mean (14.176) The table also showed that the TTMC group has the highest adjusted mean of(16.776) among the groups, followed by the CM group(14.316) while the control group has the least adjusted mean (11.556)

7. Discussion

The findings of this study revealed that the TTMC Model of Instruction produced a better performance and retention in students than the conventional model of instruction (the lecture method). As method that helps to the identification of several alternative conceptions that are held by students in science topics taught. This corroborate the findings of Lin (2004) Lin identified 14 alternative conceptions that were held by at least 10% of the 161 science majors in Year 11 ,The belief that solar energy is directly utilized for cellular activity suggest that the students probably did not possess basic understanding of the processes of photosynthesis and respiration in plant and other alternative conception identified provide useful indications for the teacher to plan for a re-teaching. The findings also supported Garnett and Hackling (9095) belief that the learning of a concept or skill depends upon the learning of the prerequisite concepts or skill. This was confirmed by the better performance and retention the TTMC Model of diagnostic produced in students exposed to it than those exposed to the conventional Model of instruction.

The administration of two-tier diagnostic instruments showed that it had the highest beta weight on students" performance in Basic Science and technology as it produces the strongest strength for predicting performance among the students. The study showed that a good teaching method increases or boosts the academic achievement, interest and attitude of students in basic science. This is supported by By Adodo and Gbore (2012) that teachers should start introducing other interesting teaching methods in the teaching of science, as this change will not only bring about an improvement in students performance in science but will also bring about lasting and permanent positive attitude towards science and that there is a reciprocal relationship between interest and learning achievement as one reinforces the other.

The TTMC Model of assessment has been found to give room for the teacher to vary his methods of teaching to achieve his objectives. It also encourages students to rely on one another as they worked in small groups on the activities more than on the teacher. These findings have implications for the design of instruction in basic science by the basic science teachers that will always probe into student's knowledge of those concepts which serve as foundation to the new concept to be taught, especially, the relevant mathematics concepts and take the pain to remedy such requisite concepts before going ahead to teach the desired concept as this will give the students a stronger footing. This is necessary as any

structure built or erected on a faulty foundation is bound to collapse. It is only when students have a meaningful understanding of what they are taught and not just rote learning, that they can rightly apply that knowledge to solve real life problems as they come their ways.

This study also reminds teachers of the strength of previous knowledge of Basic Science concepts in teaching basic science. It also shows that remedying student's weakness in Basic Science knowledge and skills enhances their performance as they now find it easier to learn basic science especially those concepts which require certain mathematical knowledge and skills for their understanding. This finding is in support of Ausubel's belief that meaningful learning occurs when there is interaction between the learners appropriate elements in the knowledge that already exists and the new materials to be learnt.

8. Conclusion and Implications for Pedagogical Practice

Since reforms in science education generally place greater emphasis on the content of curricula than on new assessment procedures. The use of two-tier diagnostic tests used in this study helped to address many   concerns about current lessons evaluation and assessment practices at its potency at overtly assessing the outcomes of students' thinking within a specified context rather than assessing knowledge of information. By using these diagnostic instruments at the beginning or on completion of a specified topic, science instructors can achieve better understanding about the nature of students understanding and the existence of any alternative conceptions are identified, in that way, science instruction can be modified to remedy the problem by developing and/or utilising alternative teaching approaches that specifically address students non-scientifically acceptable conceptions. TTMCQs Model of assessment have proved to be very successful in educational settings (Boo 2002; Boo Hong Kwen, and Kok Cheng 2005) and it has also shown in this study that not only do they help to test student understanding, but they also serve to test student higher level of cognitive thinking and help to identify misconceptions students were having   as well as increasing the students' retention of knowledge.

9. Recommendations:

TTMC is an assessment items that require an explanation or defence of the answer. The basis for this type of assessment argues that teachers must consider the intuitive knowledge base that students have already conceded before coming to the class. Teachers should try to understand students thinking of science concepts and relate their teaching to this student knowledge.

To bring these approaches to fruition by science teachers, a considerable amount of in- service education/workshop and seminars for teachers about how to develop and the use of this alternative form of lesson evaluation and assessment questions model(TTMC) will be essential and is thus recommended.

Refereces

Adodo, S O and Gbore L O (2012) Prediction of attitude and interest of science students of different ability on academic performance in Integrated Science. International Journal of Psychology and Counseling. Vol. 4(6), pp. 68-72, Kenya.

Alison Cullinane and Maeve Liston (2008): Two-Tier Multiple Choice Questions: An   Alternative Method of Formative Assessment for First Year Undergraduate Biology Students. Shannon Consortium, University of Limerick.

Boo Hong Kwen, And Kok Cheng (2005): Using two tier reflective multiple choice questions to cater to creative thinking . AARE 2005 International Education Research Conference National Institute of Education, Nanyang Technological University Singapore

Beetlestone, F. (1998). Creative children, Imaginative Teaching. Open University Press.

Black, P.& William, D (1998). Inside the black box: Raising standards through classroom assessment Phi Delta Kappan , 80 139-144..

Boo, H.K. (2002) Primary science assessment and its contribution to the nurturing of creativity. Proceeding of the International Education Research Conference organized by the AARE (Australian Association for Research in Education), 1-5 December 2002, Brisbane, Australia ISSN 1324-9320.

Boo, H.K. (2003a). Primary science assessment in the context of Thinking Schools Learning Nation vision. Teaching and Learning, Singapore, (2003), 24(2), 131-140.

Boo, H.K. (2004). The role of school science assessment in enhancing students' creativity. In A. Khoo, M.A. Heng, L. Lim, L. & R.P. Ang (Eds.), Innovation and Diversity in Education (pp. 17- 28). McGraw Hill.

Çataloglu, E. (2002). Development and Validation of an Achievement Test in Introductory   Quantum Mechanics: The Quantum Mechanics Visualization Instrument. The Pennsylvania State University.

Cengiz Tüysüz (2009): Development of two-tier diagnostic instrument and assess students' understanding in chemistry Scientific Research and Essay Vol. 4 (6) pp. 626-631.

Chandrasegaran AL, Treagust DF, Mocerino M (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. Chem.Educ. Res Practice, 8 (3): 293-307.

National Policy on Education (2004) Federal Republic on Nigeria

Griffard, P. B., & Wandersee, J. H. (2001). The Two-Tier Instrument on Photosynthesis: What Does It Diagnose? International Journal of ScienceEducation, 23 (10), 1039-52.

Haladyna, T. M. and Downing, S. M. (1989) 'A Taxonomy of Mutilple-Choice Item-Writing Rules ', Applied Measurements in Education, 2(1), 37-50.

Haslam and Treagust (1987) 'Diagnosing secondary students' misconceptions of Photosynthesis and Respiration in plants using a two-tier multiple choice instrument', Journal of Biological Education, 21(3), 203-211.

Ministry of Education Singapore (2004a). Strategies for Active and Independent Learning (SAIL). Singapore: Photoplastes Pte Ltd.

Ministry of Education Singapore (2004b). A Guide to Teaching and Learning of Primary Science.Ng, A.K (2004). Liberating the Creative Spirit in Asian Students. Pearson Prentice Hall.10

Odom, A. L., & Barrow, L. H. (1995). Development and Application of a Two-Tier Diagnostic Test Measuring College Biology Students' Understanding 89 of Diffusion and Osmosis after a Course of Instruction. Journal of Research in Science Teaching, 32 (1), 45-61.

Mann, M. and Treagust, D. F. (2000) 'An instrument to diagnose conceptions of breathing, gas exchange and Respiration ', in Paper presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans, L April 28 - May 1, 2000,18

Treagust, D. (2006) 'Diagnostic assessment in science as a means to improving teaching, learning and retention', in UniServe Science, The University of Sydney, 28 September 2006.

Tsai, C. C., & Chou, C. (2002). Diagnosing students' alternative conceptions in science. Journal of Computer Assisted Learning, 18, 157-165. Williams, J. B. (2006) 'Assertion-reason multiple-choice testing as a tool for deep learning and understanding. ', Assessment & Evaluation in Higher Education, 31(3), 287-301.

Chen CC, Lin HS, Lin ML (2002). Developing a two-tier diagnostic instrument to assess high school students' understanding, The formation of image by plane mirror. Prac. Natl. Sci. Counc. ROC (D), 12(3): 106-121

Chen CC, Lin ML (2003). Developing a two-tier diagnostic instrument to assess high school students' understanding, The 4th International Conference of the European Science Education Research Association (ESERA), Netherlans.

Chiu ML, Chiu MH, Ho CY (2002). Using Cognitive-Based Dynamic Representations to Diagnose Students' Conceptions of the Characteristics of Matter. Proc. Natl. Sci. Counc. ROC (D) 12 (3): 91- 99.

Chou CC, Chiu MH (2004). A two-tier diagnostic instrument on the molecular representations of chemistry: comparison of performance between junior high school and senior high school students in Taiwan. Paper presented at the 18th International Conference on Chemical Education, Istanbul, Turkey.

Garnett PJ, Garnett PJ, Hackling MW (1995). Students' alternative conceptions in chemistry: a review of research and implications for teaching and learning. Stud. Sci. Educ. 25: 69-95.

Kwen BH, Cheng AK (2005). Using two-tier reflective multiple choice questions to cater to creative thinking. International Education Research Conference, Singapore, AARE

Merriam SB (1998). Qualitative research and case study applications in education, Jossey-Bass Publishers, San Francisco, CA.

Novak JD (1996). Concept mapping: a tool for improving science teaching and learning, in Treagust DF, R. Duit and. Fraser BJ (Eds.), Improving teaching and learning in science and mathematics, Teachers College Press, New York, pp. 32-43.

Taber KS (2001). Building the structural conception of chemistry: some considerations from educational research. Chem. Edu. Res. Pract. 2:123-158.

Tan KCD, Treagust DF (1999). Evaluating students' understanding of chemical bonding. Sch. Sci. Rev. 81 (294): 75-83.

Tan KD, Taber K, Goh NK, Chia LS (2005). The ionization energy diagnostic instrument: a two-tier multiple-choice instrument to determine high school students' understanding of ionisation energy. Chem. Educ. Res. Pract. 6 (4): 180-197.

Treagust DF (1995). Diagnostic assessment of students' science knowledge. In: Glynn SM, Duit R (Eds.), Learning science in the schools: Research reforming practice. Mahwah, New Jersey: Lawrence Erlbaum Associates. pp. 327-346.

Treagust DF (2006). Diagnostic assessment in science as a means to improving teaching, learning and retention. UniServe Science – Symposium Proceedings: Assessment in science teaching and learning, Uniserve Science, Sydney, Australia, p. 1-9 Tuysuz 631

Tsai CC, Chou C (2002). "Diagnosing students' alternative conceptions in science through a networked two-tier test system". Int. J. Comp.Assisted Learn. 18 (2): 157-165.

Wang J R (2004). Development and Validation of a Two-Tier Instrumentto Examine Understanding of Internal Transport in Plants and theHuman Circulatory System. Int. J. Sci. Math. Edu. 2: 131-157.