TÜRK FEN EĞİTİMİ DERGİSİ
Yıl 9, Sayı 2, Haziran 2012

Journal of
TURKISH SCIENCE EDUCATION
Volume 9, Issue 2, June 2012

http://www.tused.org

Engaging Secondary Grade Physics Students in Developing Test Items

Fayyaz Ahmad FAIZE[1], Muhammad Arshad DAHAR[1]

[1] Dr., International Islamic University, Department of Education, Islamabad-PAKISTAN

Received: 05.02.2011 **Revised:** 20.07.2011 **Accepted:** 15.08.2011

The original language of the article is English (v.9, n.2, June 2012, pp.3-11)

ABSTRACT

This research aimed at investigating the effect of engaging Physics students of secondary grade in constructing test items on their test performance. Data were collected through tests. Two equivalent groups of Physics students were formed based on their performance on Physics pre-test. Both the groups were taught the same Physics content with similar teaching strategies except that the students in the experimental group were engaged in developing test items when a unit is completed. After four weeks, a post-test was administered to both groups containing factual and conceptual items in equal proportion. The data were analysed using two-factor ANOVA repeated measure design. There was found a significant difference between the mean performances of students in both the groups on the post-test as a whole. However, when the students' performance was separately analyzed on factual and conceptual items, it was found that the students in the experimental group showed better performance on factual items while, there was no significant difference in the mean score of students on conceptual items.

Keywords: Student Constructed Questions; Factual Items; Conceptual Items; Physics Test; Students' Performance.

INTRODUCTION

Test is a medium to judge students understanding of the content taught. It also provides feedback to teachers about their teaching methodology. Tests are used to measure self-esteem, aptitude, intelligence, stress, creativity, behaviour and traits (Scott & Morrison, 2005). A good test needs to be reliable and valid. Usually, it is the subject teacher who construct the test while, the students attempt it. However, can the students be assigned an active role in the learning process through engaging them to develop test item. According to some researches, when the test takers become the test maker, it makes the learning process imaginative and interesting (Iwasaki, 2008; Kaufman, 2000). But, is there any improvement in the students' performance through such practice? According to Black and William (1998), when the students assess and monitor their progress, their academic performance get improve. This kind of activity increases responsibility on the part of students for self-learning (Cyboran, 2006),

Correspondence Author email: fayyaz3shah@yahoo.com © ISSN:1304-6020

develops and refines critical thinking (Cooper, 2006) and improves academic performance (Ross, Hogaboam-Gray & Rolheiser, 2002).

Marks on students' self-assessment test will be acceptable to teacher if these marks are highly correlated to what the teacher would give on a similar test. This is possible if the marks on students' self-assessment test are not formally recorded. The students can mark their own test in the light of certain pre-established criteria or some model answer key. This activity would also spare the teacher's time for other educational activities. Usually, the teachers do not prefer students' self-assessment because the students' marks are not consistent with the teachers' marks. However, some researchers advocate that the benefits of students' generated items and assessments are so great that the teacher shall trust their students for this activity (Boud, 1989).

The students shall be properly trained to write the test items (Clarke, 1989). Such tests shall be a part of formative evaluation rather than summative evaluation and shall be used only for learning purposes (Kari, 1990). The students' made test reduces test anxiety in students, provides an opportunity to students to contribute their knowledge to the learning process and improves their test skills (Iwasaki, 2008).

However, measurement and testing in Pakistan needs to be reformed and improved. The science teachers in Pakistan teach science subjects on the same pattern as religious schools teach the theology subjects. The focus in teaching is on the memorization of the concepts and figures rather than on understanding. The memorized knowledge is then produced in the examination. The students have no opportunity to test even the knowledge they have memorized (Sadiq, 2003)

Regarding secondary classes, the tests are made and checked by the subject teachers except the terminal examination which is prepared by an external examination board. These exams are taken at grade 10 after which the students apply for admission in college. Meanwhile, the secondary level teachers do not make the test items from themselves. The teachers copy the questions from the past papers and then give it to the class for assessment purpose. The test items are mostly focus on memorization. The examination papers keep on repeating a few long questions in a periodic manner. Even, a guess paper is sometime available which lists the expected questions that might come in the external examination.

In order to improve the students' performance at the secondary level, the science teachers do some practice with the past papers questions so that the students may have information on how they would be assessed on a particular content. However, there is little concept of involving students in writing test items. The researcher aimed to involve Physics students in writing test items. The objective of this involvement was to observe the effect of this involvement on students' test performance. The test items may evaluate different assessment objectives. In this study, the researcher has delimited the assessment objectives to recalling and understanding component.

The test items used in the research were multiple choice type. According to Faize, Dahar and Niwaz (2010), multiple choice items is the most widely used format for assessment purposes in majority of the fields at present however; it is a also source of anxiety to students. The researcher chose the multiple choice format to help students reduce anxiety by involving them in developing such items. The students developed two types of items; factual and conceptual. The items that assess the students on some recalling component was termed 'factual items' and items that require the students to analyze and synthesize the acquired knowledge to answer a question was termed 'conceptual items' in this paper. The researcher aimed to explore whether involving Physics students in developing test items as a class activity can improve their performance on an actual test. Further, the researcher also aimed to find out the performance of Physics students on factual and conceptual items separately and explore if there exist any significance difference in the mean score of students on factual and conceptual items.

Review Of Literature

The activity of engaging students to participate in the evaluation process proved beneficial to various researchers. Foos (1989) conducted a research on 94 students enrolled in introductory psychology class. Two groups were formed. Half of the students were asked to write multiple choice questions (MCQs) and the half to write essay type questions. The students were encouraged to write factual as well as conceptual items. After practice with writing test questions, a teacher-made test was administered to the students. The teacher ensured that the student-written items were not included in the test. It was found that there was no significant difference between students' score who wrote MCQs and those who wrote essay type questions. Students who wrote questions performed significantly better than students who did not practice with writing test questions. Students who wrote essay questions did not perform significantly better on essay type questions than students who wrote MCQs and conversely. However, it appears very surprising that despite the practice with writing essay type questions, the students did not perform significantly better on essay items and the same result was obtained for students who practiced with writing MCQs.

Few years later, Kerkman, Kellison, Piñion, Schmidt & Lewis (1994) investigated a similar relation between students' developing test items and its effect on their test performance. They divided the students into two groups- the experimental and the control group. Both the groups experienced the same lectures and class activities with the exception that the experimental group was asked to write MCQs on daily basis after their reading assignment. The students were told to avoid questions that require names or dates as their answer. The students received marks for writing test items that fulfil certain pre-set criteria. The teacher would choose the best students made questions and would administer it as a surprise test to the students. The test was shuffled and handed over to other students for marking purpose. The control group would receive a comparatively easy test to compensate for not writing test items. The researchers found that the students in experimental group scored significantly higher than the control group. Thus, writing test questions improved students' performance. However, it was not made clear in the above-mentioned research what was the pre-set criteria that a test item had to fulfil. Secondly, the students generated test items were administered to the same students as a surprise test. It is understood that the students would perform better on such a test that contains the items developed by them.

In order to include teacher's made test items instead of students' made items, a better strategy was adopted by Kaufman (2000). He also formed smaller groups containing three or four students. Each group was asked to prepare a lesson and then present it in the class. The students were instructed to write test items related to the lesson. The teacher then administered a test to the students. This test comprised of students generated items as well as teacher made items in equal proportion. After the teacher marked the test, the students were asked to correct their mistakes. The students would seek advice from other groups as well in order to understand the questions prepared by that group. Thus, the students acted as test-maker, test-taker as well as peer-tutor. The students were enthusiastic and helped one another to learn and understand the content in a better manner. Writing test items helped the students in understanding the material better. However, Kaufman also included the student generated items in the test and thus the student's performance was questionable. Moreover, the researcher did not give any statistical evidence of his findings.

Regarding students' attitude towards writing test items, Giles et al (2004) found that the students involved in preparing test items became very creative and competent in fulfilling the task. The students also enjoyed the process and regarded it as a useful activity and valuable experience in the field of testing and evaluation.

The research literature cited above was conducted in the humanities subjects and the respondents were college students. The purpose of the present study was to understand whether engaging physics students in developing test items has a positive impact on their test performance. How will the students of secondary grade perform on factual and conceptual items after practising with writing test items? With these objectives, the researcher conducted an experimental study to investigate the effect of engaging Physics students in developing test items on their test performance.

METHODOLOGY

a) Research Design

The study was experimental and the researcher used pre-test post-test equivalent group design.

b) Sample

The researcher selected a government comprehensive school in the federal territory of Islamabad where the required number of Physics students was available and the administration gave consent for conducting the research. The sample comprised of 82 physics students of grade 10 registered for the session 2010-2011.

c) Developing Instruments and Pilot Study

In order to observe the effect of developing test items by Physics students on their test performance, the researcher collected data through two Physics tests called Physics test 1 and Physics test 2. The first test, Physics test 1 was administered before the formation of groups while, the second test, Physics test 2 was administered after the treatment period of four weeks. Both the Physics tests were prepared by the researchers with the help of two Physics teachers having more than 10 years of teaching experience. The items in the tests were multiple-choice type with four distracters. There were 50 MCQs in both the tests. Three physics experts validated the tests for face and content validity. The test was piloted on twenty Physics students in the same school. These twenty students were not included in the sample from which the data were collected later. The students' performance on the pilot test was analysed using split-half reliability for internal consistency. The reliability coefficient on the pilot test was 0.86. The wording of the some items was also improved based on the suggestions and difficulty faced by the students. All the test items were teacher made and included factual and conceptual items in equal proportion. The MCQs were constructed according to the suggestions given by Thorndike, Cunningham, Thorndike and Hagen (2001).

d) Procedure

After administering Physics test 1 to the sample comprising 82 students, the marks obtained by the students were arranged in descending order and based on these scores; two equivalent groups of students were formed. After the formation of groups, one of the groups was randomly assigned as experimental group and the other control group (See fig. 1).

There were forty one students in each group. However, one student from each group was later dropped due to short attendance. Both the groups were taught separately by the same teacher. The time of the class of each group was regularly interchanged to control the time variable. Both the groups were taught the same units. These units include temperature, heat capacity and transfer of thermal energy related to thermal Physics section of grade X syllabus. However, the students in the experimental group were involved in an extra activity to develop test items related to the content taught. The students were asked to write factual as well as conceptual items in equal proportion. The students were given clear instructions and guidelines on how to write a factual or a conceptual item. For this purpose, the teacher presented different examples of test items to the students and guided them about its construction. The researchers also provided the students a copy of suggestions for writing multiple choice items taken from the book of Thorndike et al. (2001). Below is given an example of factual and conceptual item.

Table 1. *Factual and conceptual test Item*

Factual item	**Conceptual Item**
The lower fixed point in calibrating a thermometer is A. -273 ^{0}C B. 0 ^{0}C C. 4 ^{0}C D. 100 ^{0}C	A liquid-in-glass thermometer is used to measure the boiling point of water. The liquid in the thermometer is replaced by another liquid which expands more for the same temperature rise. The new thermometer will have A greater sensitivity and greater range. B greater sensitivity but less range. C the same sensitivity and the same range. D the same sensitivity but greater range. (UCLES 2004, 5054/01/M/J/04/ Q17)

The students worked in small groups of 3 to 4. The questions developed were multiple-choice type. The students were encouraged to discuss the items in their group. The teacher facilitated the students and provided necessary help to correct the wording of the question. Two teacher assistants also helped the physics teacher during the lesson so that the group is not distracted while developing test items. The teacher assistants would also check the constructed items and would help the group in refining and improving the wording of the items. The teacher assistants possessed a graduate degree in science and a Bachelor degree in education and three years of teaching secondary students. During the time, when the experimental group was developing test items, the control group was given teacher-made test to compensate for the time. The two teacher assistants would check the students work and would provide necessary help to the students in answering an item correctly. At the end of fourth week, the Physics test 2 was administered to both the groups as a surprise test.

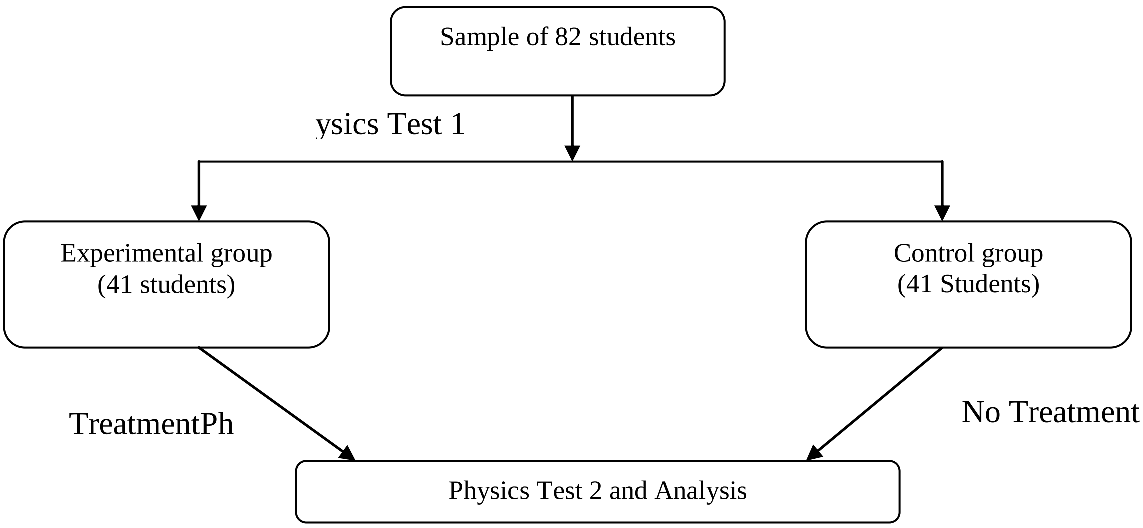

Figure 1. *Design of group formation*

FINDINGS

The Physics test 2 was marked by the concerned Physics teacher according to a pre-made marking criterion to ensure reliability in scores. The score of each student on factual and conceptual items was separately entered in the SPSS 18.0 data file for both the groups. The sum of score on factual and conceptual items gave the total score obtained by a student. The mean marks and the sum of squared deviation in each group were calculated. Finally, the researcher calculated the F-value using two way ANOVA repeated measure design to observe if there exists any significant difference in mean score of experimental and controlled group on factual and conceptual items.

Table 2. *Group wise Comparisons*

(E) Group	(C) Group	Mean Difference (E-C)	Std. Error	Sig.
experimental group	Controlled group	5.30*	.86	.000

The researchers calculated mean score in both the groups to observe the mean performance of students. It was observed that the students in experimental group have higher mean score as compared to control group on the Physics test 2 as a whole. The difference in the mean score of the two groups was 5.30 which was significant at 0.05 level (Table 2). This was expected as was found in the previous studies. Thus, the treatment has an effect and the practice in writing test items by the students improved their test performance as a whole. This was also tested using two-way ANOVA repeated measure design giving a similar result; $F(1, 78) = 324.63$, $p < .05$, $\eta^2 = 0.81$ (Table 3)

Table 3. *Students' overall performance on the post-test as a whole*

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Test item	8643.60	1	8643.60	324.63	.00	.81
Error(Test item)	2076.80	78	26.62			

As the Physics test 2 was comprised of two sections, the factual test items and the conceptual test items; the score of students in the experimental group and control group was analyzed on both the test items. The data in table 4 revealed that the students in both the groups performed higher on factual items as compared to conceptual items. The overall mean score on factual items was 33.30 against 18.60 for conceptual items.

Table 4. *Mean score of students in both the groups*

	Groups	Mean	Std. Deviation	N
Score on factual items	**experimental group**	37.80	7.23	40
	controlled group	28.80	3.41	40
	Total	33.30	7.21	80
Score on conceptual items	**experimental group**	19.40	5.82	40
	controlled group	17.80	3.78	40
	Total	18.60	4.94	80

It was also observed that the students scored higher on factual items as compared to conceptual items having mean score of 33.30 and 18.60 respectively. The students in the control group managed to get high score on factual items even without developing test items.

The next matter of interest with the researcher was to observe and analyze the performance of the students in both the groups on the test items separately. The objective was to observe the performance of students in comparison to item type.

Table 5. *Students score on factual items*

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Contrast	1620.00	1	1620.00	50.69	.000	.39
Error	2492.80	78	31.95			

Firstly, the score obtained by students in both groups on factual items was calculated and then analysed. It was found that there was significant difference in the mean score in the experimental and control group on factual items. $F (1, 78) = 50.69$, $p< .05$, $\eta^2 = 0.39$ (Table 5). Though, the students in both the groups performed higher on factual items, however, the difference in the score was still significant. Thus, engaging the Physics students in developing factual test items improved their performance on a teacher's made tests. The reason for this may be that, while writing factual items, the students in the experimental group got the opportunity to observe and remember the answer, which the students recalled successfully on a teacher made test.

Table 6. *Students score on conceptual items*

Source	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Contrast	51.20	1	51.20	2.12	.15	.027
Error	1880.00	78	24.10			

Another interesting finding was the comparative performance of students in both groups on conceptual items. The students in the experimental and control group performed lower on conceptual items. When analysed statistically, the researcher found no significant difference in the mean performance of students in both the groups on conceptual items. $F (1, 78) = 2.12$, $p> .05$, $\eta^2 = 0.027$ (Table 6)

It was harder for the students to score on conceptual items as compared to factual items. This kind of result was not against expectation. The students in Pakistan do better on factual items as compared to conceptual items. The reason for this might be the examination system, which put more emphasis on rote learning, and the students found it easier to learn material requiring memorization. The teaching practices are also geared around factual information with less emphasis on understanding and application of concept.

Consequently, despite involving the experimental group in writing conceptual items, there was no significant difference in the mean score of both the groups. The reason for this may also be the broad scope of conceptual test items. The construction of conceptual test items involves high cognitive and analytical skills; it was difficult for the students to score on these items. Moreover, a conceptual item may ask about a concept in a number of ways therefore, it might have been difficult for the students to answer on an item.

CONCLUSION and RECOMMENDATION

Writing test items is a teacher's job. However, the students can be engaged in writing test items. An investigation in this regard found that the students in the experimental group performed significantly better than the students in the controlled group. Developing test items improved students' performance on the post-test as a whole. Analysing students' performance in the two groups on factual items, it was observed that both groups scored higher on factual component of the post test. However, the two groups still differed significantly with experimental group performing higher on factual items. Comparing the two groups on conceptual items, the overall mean was lower than the mean score for factual items. However, the two groups still differed significantly on conceptual items. The lower performance of the two groups on conceptual items requires more research. May be the students were weak in writing conceptual test items in physics. The possible cause for this may be the testing system, which is more focussed on rote learning and which stops mental growth and blocks innovative learning (National Education Policy, 2009). A teacher being qualified and with knowledge of past papers questions can develop conceptual items, however, science students have a limited knowledge of past papers questions and assessment skills. Thus, the conceptual items were difficult to construct and scored. Improving students' performance on conceptual part requires more research in testing and evaluation techniques. Writing conceptual items demands more understanding of the content, practice with the past papers questions and knowledge application. The need is to shift from rote learning to understanding and application of concepts. Though, the examination boards have introduced some changes in the paper pattern in the last few years, and are including questions assessing understanding and application of concepts. However, the science teachers still use the lecture method for teaching science and are not competent in teaching through activity oriented method in Pakistan (Faize & Dahar, 2011). As the science teachers would come out of the traditional teaching methods and modify their teaching styles, the cognitive skills may improve which may result in improving students' performance on conceptual items as well. A research on investigating the effect of developing test items on the students' performance at University level may form an interesting area for further research. Similarly, research can also be conducted in other science subjects at secondary or higher level for understanding the performance level on conceptual items.

REFERENCES

Black, P., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80*(2)*, 139–148.

Boud, D. (1989). The role of self-assessment in student grading. *Assessment & Evaluation in Higher Education*, 1469-297X, 14*(1)*. 20 – 30

Clark, D.F. (1989). Materials adaption: Why leave it all to the teacher? *ELT Journal*, 34*(2)*, 133-141

Cooper, D. (2006). Collaborating with students in the assessment process. *Orbit*, 36(*2*), 20–23.

Cyboran, V. (2006). Self-assessment: Grading or knowing? *Academic Exchange Quarterly*, 10(*3*), 183–186.

Faize, F. A. & Dahar, M. A. (2011). Physics Teachers Teaching Competency in Pakistan- Perception of 'O' Level Students. *American Journal of Scientific Research*, 16. (in press)

Faize, F. A., Dahar, M. A. & Niwaz, A. (2010). Improving Students Performance on Multiple-Choice Questions in Physics- Practice in Justifying the Options. *International Journal of Academic Research*, 2(6), 204-207

Foos, P. W. (1989). Effects of student-written questions on student test performance. *Teaching of Psychology*, 16, 77-78.

Giles, A., Martin S., Bryce, D. &Hendry G. (2004). Students as partners in evaluation: student and teacher perspectives. *Assessment & Evaluation in Higher Education*. 29 (6), 682-685.

Iwasaki, I. (2008). Using Student-made Quizzes in the EFL Classroom. *The Internet TESL Journal*, Vol. XIV (12). http://iteslj.org/Techniques/Iwasaki-Student-madeQuizzes.htm

Kari, S. (1990). Let your students write their own tests. Paper presented at the annual meeting of the International Association of Teachers of English as a Foreign Language, Dublin: Ireland. (ERIC Document Reproduction Service No. ED 322 736)

Kaufman, L.M. (2000), Student-Written Tests: An Effective Twist in Teaching Language. *The Journal of the Imagination in Language Learning and Teaching*, 5. http://www.njcu.edu/cill/vol5/kaufman.html

Kerkman, D. D., Kellison, K. L., Piñion, M. E., Schmidt, D. & Lewis, S. (1994). The quiz game: Writing and explaining questions improve quiz scores. *Teaching of Psychology*, 21(2), 104-106. doi:10.1207/s15328023top2102_11

National Education Policy (2009). Ministry of Education, Government of Pakistan.

Ross, J. A., Hogaboam-Gray, A., & Rolheiser, C. (2002). Student self-evaluation in grade 5–6 mathematics: Effects on problem solving achievement. *Educational Assessment*, 8(1), 43–58.

Sadiq, A. (2003). Evaluation of Scientific Enterprise in Pakistan. In: Dr. Inayatullah (ed.), *Towards Understanding the State of Science in Pakistan*. Karachi, Muizz Process.

Scott, D. & Morrison, M. (2005). *Key ideas in educational research*. London: Continuum International Publishing Group.

Thorndike, R. M., Cunningham,, G. K., Thorndike, R. L., & Hagen, E. P. (2001). *Measurement and Evaluation in Psychology and Education*, 5th edn., New York: Macmillan Publishing Company.

TÜRK FEN EĞİTİMİ DERGİSİ
Yıl 9, Sayı 2, Haziran 2012

Journal of
TURKISH SCIENCE EDUCATION
Volume 9, Issue 2, June 2012

http://www.tused.org

Gender-Based Differences in Language Learning Strategies of Science Students

M. Naci KAYAOĞLU[1]

[1] Assist. Prof. Dr., Karadeniz Technical University, Faculty of Letters, Trabzon-TURKEY

Received: 22.02.2011 **Revised:** 06.02.2012 **Accepted:** 25.04.2012

The original language of the article is English (v.9, n.2, June 2012, pp.12-24)

ABSTRACT

Since the early 1970s there has been considerable research interest in the strategies that foreign or second language learners use in learning and using the target language. While the accumulated literature to date suggests a possible relationship between strategy use and foreign language learning success, a substantial number of studies have been undertaken in relation to various factors which have been found to affect strategy choice and use. Although gender is typically a significant factor in other disciplines in particular in educational and psychological research, gender has only received sporadic attention in the studies of language learning strategies. This study, therefore, is an attempt to explore the influence of gender on language strategy use.

A total of 115 science students were involved in this study. Survey method was used as data collection instrument. The results indicate that female students reported higher strategy use. More specifically, female learners showed greater use of the five major strategy categories (memory, compensation, cognitive, metacognitive and social strategy categories).

Keywords: Gender Differences; Science Students; Language Strategies.

INTRODUCTION

Learning strategies have become widely recognized as the central element in several models of language learning. The increasing awareness of the behaviors which learners consciously or unconsciously employ while learning a foreign language has been probably one of the most important outcomes of the movement in its transition towards a learner-centered approach to language learning. This change has been reflected in various ways in language education, ranging from the instructional materials to the curriculum characterized as the learner-centered curriculum (Nunan, 1988).

It is obvious that this growing recognition came with the emergence of cognitive psychology within which learners are seen to be actively involved in the process of learning. It is in direct response to the behaviorism theory of language learning which dominated the

Correspondence Author email: naci@ktu.edu.tr

© ISSN:1304-6020