© 2016 IJSRSET | Volume 2 | Issue 5 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section: Science and Technology

How Does Item Difficulty Justify Learners' Poor Attendances in ELT Context?

Behnam Behforouz
Foundation English Department, University of Buraimi, Oman

ABSTRACT

This study investigated the relation between degree of test difficulty and students' performances in midterm and final exams. For this purpose, 24 Arab students were selected. Examination of results revealed that of course there is a relation between these two variables but it is not a big gap.
Keywords: Item Difficulty, Poor Attendance, ELT Context

I. INTRODUCTION

Regarding previous researches about students' achievement (Qayyum& Ismail, [32], Hong, [17], Nami, [29]), there are many factors which affect students' performance in a test. Item difficulty, motivation, time, etc., all are among those factors. According to Marso [24], educators always deal with the problem of classroom tests. Some like believes that when the item difficulty reaches 50 percent, the test has maximum reliability and validity, but classroom activists believe that difficult tests decrease students` motivation leading to poor achievement (Wood, 1961). Expected test trouble is generally recognized as a student formulated model and has been found to have significant effect on test performance. (Marso, [24] cited in Qayyum & Ismail, [32]).

Making reasonable and orderly assessments of others execution can be a challenging task. Judgments' can`t be made exclusively on the premise of instinct and heedless speculation, or custom (Sax, 1989). Educators, managers, and others in evaluative positions utilize an assortment of instruments to help them in their assessments. Tests are instruments that are frequently used to encourage the assessment procedure. To develop the norm-referenced tests for instructional purposes, to evaluate the impacts of educational plans, or to investigate examination purposes, it can be vital to conduct item and examine the analysis. (Matlock-Hetzel, [25]).

1.1. Statement of the problem

State-administered tests (i.e. estimation tools that are utilized to assess levels of particular proficiencies, aptitudes or abilities, for example, knowledge of science) are utilized as a part of instruction for a few purposes, including outlining and measuring changes in understudy accomplishment, and in addition for symptomatic and aptitude purposes (Shephard, 2001; Black and Wiliam, 1998). At the point when tests are likewise utilized as a feature of the premise for granting understudy grades, they can contribute towards evaluating which schools the participants may go to later on, the understudies' self-viability (Bandura, 1997).

Suppose you were sitting, tensely waiting for an exam to start. What might you do? Would you utilize your additional time before exam to concentrate on? How might you contemplate? What system would you utilize? Foos [12] suggested some of these questions in a study where students were informed that an exam they were concerning to take would be either simple or hard. All participants were given fifteen minutes to concentrate on. Those who thought the test would be hard, performed superior to those participants who thought the test would be simple. He [12] reasoned that students who thought the exam would be troublesome

were spurred to study and to work harder than others who suspected the test would be simple.

According to the above statement, it seems that whenever students are ready for difficult exams, their performance would be better than the time when they think that the exam is easy. It's because during hard exams students take it seriously, think more, study hard, investigate all alternatives, while in opposite side it never happens, in easy exams. The students suppose that they pass the test easily because they were told. In fact, the test is not easy but its psychological effect of the speech on students.

During a semester, teachers work hard, cover different materials, supply more formats of tests for the students. They think that final exam may be easier for their students, for this reason all the quizzes are difficult to some extend and the logic of such teacher is to make his/her students ready for the final exam. But the problem is that sometimes wrong professionals are selected to design questions (people with no EFL background), in this situation they make very hard and challenging tests without caring students level and knowledge. I believe that in these cases, students motivation goes down, they lose their confidence, and they fail certain and very important items during the exam session. Here I think the attempts of a teacher in whole semester ruin.

The purpose of this study is to investigate students' performance during a semester on some difficult tests, comparing the results of their performance in midterm exam which was easier than final exam. I want to investigate and explain why students did very bad in final exam while they received highest marks during class quizzes and activities.

1.2. Research question

The question addressed in this research is 'Does the difficulty of an exam support student's poor performance in final evaluation? '

1.3. Research hypothesis

Regarding research question above the following hypothesis is formulated:
H: Item difficulty has direct and negative effect on student's performance.

II. THEORETICAL BACKGROUNDS

2.1 Test Anxiety

Anxiety s is normal among pupils at all levels of instruction. A particular kind of tension experienced amid testing circumstances, test nervousness (TA), has been appeared to inspire passionate reactions that add to errand impedance and lower scholarly testing execution (Sarason, [38]). Test nervousness is additionally identified with fears of negative assessment, the detesting of tests, and poor study propensities (Hembree, [16]). Moreover, abnormal amounts of test uneasiness have been appeared to be adversely associated to numerous variables, for example, IQ, scholarly accomplishment, memory, and cumulative evaluation point normal (Cassady and Johnson, 2002; Hembree, [16]). In a meta-examination of 562 studies on test nervousness in scholastic settings led by Hembree [16], results found that test uneasiness is steady in decreasing execution.

Different zones of cognition are additionally contrarily affected by test tension. For example, stressing is said to utilize a portion of the subjective preparing assets, like consideration, mindfulness, and working memory that could somehow be connected toward the testing circumstance (Wine, 1971; Lehto, 1996; Dobson and Markham, 2001). Stress and emotionality are additionally generally expected to prompt learning shortfalls (Cassady and Johnson, 2002). In particular, these learning deficiencies regularly appear in testing circumstances. As the requests of a testing circumstance expand, so does the anxiety as well (Meijer, 1996). It has been found that the discernment that people will be judged on their execution is sufficient to evoke increments in the tension they encounter (Sarason, [38]; Meijer, [26]; Foos and Fisher, [13]; Jensen-Campbell, et al., 2002). In this manner, the view of the classroom or testing environment can impact execution (Lidz and Elliott, 2000; Nelson and Knight, 2010).

This decrease in execution can be ascribed to numerous elements that impact test uneasiness. Students who stress over tests will probably feel less self-regard, take more time to finish tasks, invest more energy considering, have higher state nervousness, and expect lower achievement (Alansari, 2004;Hembree, [16]; Foos and Fisher, [13]). Test nervousness is experienced

at different degrees, be that as it may. Females tend to report more elevated amounts of test nervousness than their male partners, yet they perform similarly well on psychological measures (Hembree, [16]).

Still, a few learners demonstrate fundamentally more test tension than others. Concentrates, for example, Hancock [15] and Zatz and Chassin (1985) have split learners into groups, for example, high test-on anxious (HTA) and low test-anxious (LTA), in light of their scores on test tension measures and afterward tried them on various variables to gage execution. learners high in test tension commit more psychological assets to errand crippling contemplations, which prompts poor execution, as appeared by Zatz and Chassin (1985).

At the point when danger is brought into the condition, contingent upon the individual, the uneasiness state can be more negative or valuable to execution. Evaluative risk can be seen as a situational weight in which people accept that their execution will be contrasted with others (Zatz and Chassin, 1985; Hancock, [15]). Very stressful learners tend to score better without evaluative weight, while others with low nervousness perform better with assessment (Hembree, [16]). The profoundly restless students see the risk adversely, stressing over their capacities and performing wastefully, while the low nervousness pupils view assessment as an approach to showcase their abilities (Cassady and Johnson, 2002). In this way, keeping in mind the end goal to boost execution, the objective must be to figure out how to diminish the uneasiness while within the sight of an evaluative circumstance so learners can perform getting it done in spite of outer weights.

2.2 Learning Potential and Dynamic Assessment

On the off chance that the motivation behind regulating an allot is to locate the genuine abilities of the member, evaluations ought to minimize the impact of incidental variables (Lidz and Elliott, 2000). In this outlook, the idea of learning potential (LP) has been proposed as a technique for measuring one's ability for learning under ideal conditions.

One approach to operationalize LP is the level of information that can be accomplished in the wake of accepting help (Meijer, 2001). A more extensive definition is that LP is basically the capacity to get and apply psychological abilities (Rempfer, Hamera, Brown, and Bothwell, [33]). In this manner, LP is frequently measured utilizing dynamic assessment (DA) strategies, which implant instructions in the assessment strategy. Meijer [26] noticed that LP evaluation is utilized as a part of a wide range of settings, for example, instructive testing, genuine emotional sickness (SMI) assessment, and professional testing. The supposition is that LP assessments are more prescient of an individual's capacities than customary tests, because of the more noteworthy spotlight on the learning procedure itself rather than the deciding results. Moreover, LP evaluation advantages the individuals who are defenseless against misconception the guidelines or desires of conventional estimations because of newness, social disservices, or other learning challenges (Barr and Samuels, 1988; Glutting and McDermott, 1990; Lauchlan and Elliott, 2001).

The test then is to make a testing circumstance in which impeded students can exhibit their capacity to apply abilities alongside adapting new ones (Lidz and Elliott, 2000; Kozulin, 2005; Kozulin, 2010). DA empowers the investigation of how contrasts in individual test execution are influenced by changing the testing conditions (Bethge, Carlson, and Wiedl, 1982). Dynamic evaluation measures comprise of a pretest-guideline posttest group, with people will apply the strategies or directions learned in the second stage amid the posttest, giving scientists a superior representation of learning capacity. As a representation of DA, Wiedl [52] utilized these strategies to analyze learning potential in individuals with schizophrenia.

Wiedl [52] grouped people in light of their scores on a measure of official working (higher-request subjective capacities incorporate particular consideration, arranging, and the control of data in critical thinking). He built up a dynamic convention for the Wisconsin Card Sorting Test (WCST; Kongs, Thompson, Iverson, and Heaton, 2000), a measure of official capacity utilizing cards. In this testing organize, the WCST is managed in three trials: trial one and trial three are customary trials, controlled under the standard testing convention; trial two is an instructional trial, where the researcher gives criticism in light of the reactions of the participant. This case of dynamic assessment highlights the significance of the procedure, more than the item of perception. By supporting examinees amid a testing circumstance, the weaknesses that they experience, for

example, nervousness, ought to be to some degree mitigated, giving a more exact estimation of execution (Meijer, [26]).

One use of LP strategies has been to anxious test-takers, who may profit by the non-conventional testing arrangement of DA. It has been found that subjects have been less on edge in the wake of learning potential tests and that these learning tests are less one-sided against anxious subjects than conventional tests (Meijer, [27]). Measuring the execution of participants when help be gotten observed to be a more legitimate indicator of future execution for HTA pupils than for LTA ones (Meijer, [26]).

In addition to learners' expectations that a test may be hard, test uneasiness has additionally been appeared to affect test execution (Seipp, [43]). Actually, Hong [18] observed that uneasiness over taking a test intervened the relationship between saw test trouble and execution. That is, pupils who trusted that a test would be hard experienced more stress and thus, performed more terrible. These outcomes seem, by all accounts, to be conflicting with the Foos [12] study. In any case, test takers in Hong's [18] study were understudies selected in a measurements course and were taking a real exam for the course. It is conceivable that participant tension in this study was identified with the measure of readiness for the exam, which may likewise have been identified with exam execution. For instance, an understudy who had not contemplated for the test may have high tension and a low execution because of absence of arrangement.

To further address the perplexing relationship amongst nervousness and saw trouble on test execution, Weber and Bizer (2006) cited in [32] recommended that the sum and sort of uneasiness experienced by test-takers might be useful now and again and impeding in others, contingent upon desires that the test would be simple or hard. In Weber and Bizer's (2006) cited in [32] study, members' nervousness levels were evaluated and afterward members were given an exam with directions showing that the test they were going to take would be either simple or hard. Test execution associated with nervousness levels such that low levels of uneasiness encouraged better execution when members trusted the test would be hard and high tension brought about poorer execution in the hard test condition. The study demonstrated that foreseen test trouble did not have a reliable relationship to test execution. Weber and Bizer's (2006) cited in [32] study left open the subject of whether understudies will probably study harder for a test they accept will be troublesome when contrasted with a test they accept will be simple.

In a study, Combs, Michael, Fiore, and Poling [6] found that members who were in a classroom simulated group context, concentrated longer than others in a single testing group, particularly when they expected that the test would be hard. Moreover, participants who were in the classroom stimulus group performed more regrettable on the test than others who were in a single testing group. Though members who reported large amounts of uneasiness, as measured by the TAS, performed more terrible on the test than participants who reported lower levels of nervousness, it doesn't create the impression that tension was the intervening variable in the poorer execution of those in the group context. That is, there were no distinctions in uneasiness scores for participants in the group versus singular context.

At in the first place, it appears to be confusing that pupils in the group context would concentrate longer but perform more terrible on the test. Moreover, this is not by any stretch of the imagination astonishing. The nearness of associates with regards to the classroom likely sufficiently made social impact that members felt constrained to use whatever study time was made accessible to them. It appears pupils in this study occupied with a type of open congruity (Brehm, Kassin and Fein, [4]) in which they demonstrated their conduct after what they watched their colleagues doing, yet they might not have had the interior inspiration to ponder in a quality manner.

The lower test scores got by the participants in the group setting may be clarified by low quality study strategies joined with the excitement brought on by having others present in the room amid the testing. This impact, known as social assistance is entrenched in social mental writing (Zajonc, [54]). If participants in the group setting were fledglings in psychology, then social assistance would foresee that their execution would be ruined by the nearness of others. It is sensible to accept that they were beginners in psychology since they were selected from General Psychology (PSY 101) amid the main week of the term.

Dissimilar to past examination, this study did not reproduce past studies that have discovered foreseen trouble results in higher test scores (e.g., Foos, [12]; Weber and Bizer, (2006) cited in [32]. However, this study replicated discoveries with test uneasiness. To be specific, higher uneasiness results in lower test scores (Hong, [18]). In this study, test-takers did not demonstrate larger amounts of uneasiness as a result of the setting (group versus individual) or as a result of being informed that the test was simple or hard. Given that uneasiness levels did not rely on upon particular test-related variables, it might be sensible to estimate that the kind of tension measured here was quality nervousness, instead of state tension. In a testing connection, attribute tension is a general worry with all examination circumstances while state uneasiness is a level of stress coordinated toward a particular exam or exam setting (Hong and Karstensson, [19]). As measured by the TAS, test uneasiness mirrors a broader build steady with attribute tension. This is predictable with Weber and Bizer's (2006) cited in [32] research in that they discovered impacts with quality uneasiness, however not state tension. (Combs, Michael, Fiore, &Poling, [6])

In another study Kale, Fowler, and Rempfer (2012), inspected the impacts of students' test anxiety on their intellectual execution. Participants reporting high and low test tension were presented to two sorts of assessment – high danger and low risk. After the tests were regulated, the students' execution levels were measured.

This study found that students in the high evaluative condition did not play out another way than those who were in the low evaluative condition. This shows that all participants, notwithstanding the evaluative condition, could perform at the same level of working. An intriguing finding was that the stress part of the TAI was observed to having a negative impact on LNS execution. This outcome proposes that stressing may obstruct the working memory capacity, which ought to be inspected in future exploration. As showed by Meijer [26], the higher the requests of a testing circumstance, the more unsafe the impacts of uneasiness can be to the general execution of the individual, disturbing concentration and taking up a more noteworthy measure of the handling limit that could be utilized for other psychological procedures.

Likewise, in these information, participants reporting high uneasiness did not play out any other way than those reporting low nervousness. These discoveries are as opposed to past examination, which demonstrated that participants who report higher tension ought to perform more inadequately on psychological measures than those who report lower nervousness (Hembree, [16]). This exhibits the majority of the students, in spite of the nervousness reported, could perform at the same level of working. In any case, tension was inclining towards affecting students' learning potential, as showed by the element WCST. This recommends tension could be a variable that impacts students' capacity to learn and ought to be inspected further in future exploration.

Further, it is conceivable that other DA measures could be more profitable for this population than the WCST (Lauchlan and Elliott, 2001), which may demonstrate more illustrative of learning potential in this population.

2.3 Test Difficulty

In spite of the fact that exploration into the impacts of expected test trouble on test execution is copious, generally couple of examinations have specifically centered around prompt cautioning of test trouble. A progression of exploration on test execution utilize the approach where the notices of test trouble are given well before the exam with the goal that participants have satisfactory time to prepare. Unfortunately, the develop of cautioning students just minutes before the exam keeps on being an understudied range. Lately, Weber and Bizer (2006) cited in [32] examined the impacts of quick admonishing of test trouble on test execution in a research center setting, where participants' execution did not have any ramifications on their evaluations. The participants included sixty-two Eastern Illinois University psychology students, who were subjected to Graduate Record Examination (GRE). Despite the fact that, the significance of this exam was apparently pushed upon students, clearly their execution made little difference to their future scholarly life.

The discoveries uncovered that students with low characteristic nervousness performed better when informed that the test would be troublesome than when informed that the test would be simple. Then again, participants with high attribute tension were

defenseless to unfavorable execution when cautioned that the test would be troublesome than they did when told the test would be simple. Henceforth Weber and Bizer (2006) cited in [32] likewise discovered generous experimental backing for a curvilinear relationship between test tension and test execution.

Thiede [47] calls attention to that students were initially presented to either a more troublesome or a less troublesome test before being eventually given a last test. It was accepted that being given a test of a specific level of trouble would lead these students to foresee a last test of comparative trouble level. The variation scores of tests were then contrasted to decide the impact of expected test trouble on test execution. The discoveries of these analyses mirrored that test execution was better for students who had expected a more troublesome test when contrasted with the individuals who had foreseen a less troublesome one. The scientists referred to three primary purposes behind this: (1) an uplifted inspiration to study, (2) more noteworthy time spent on assignment, and (3) longer study time. It must be noticed that previous studies were commonly led in a setting where understudies had adequate time to ponder for the future test, which they expected to be troublesome. however, on occasion students find out about the trouble of a test only minutes preceding the test. For instance, they may become acquainted with it through other batch-mates or the instructor may drop the indication while disseminating the test paper. In this occurrence, students have no opportunity to think about the test, rendering all the aforementioned hypothetical clarifications inapplicable. Analysts trust that for this situation, nervousness will be activated. Consequently, the term 'quick cautioning' in the present study situation alludes to the notice that is given only minutes before managing the test. (Sax, Eilenberg, and Klockars, [41], Foos, [12], Eisenberger, [9], Kellas and Butterfield, [20], Weber and Bizer, 2006 cited in [32])

2.4 Role of Test Anxiety in Influencing Test Performance

At the point when students are cautioned about the trouble of a test, test nervousness is stirred above. Since test nervousness is one of the hidden builds in the present study, it is basic to comprehend it further. Test uneasiness is a broadly examined region in educational psychology and has been the subject of much insightful debate. One all around acknowledged meaning of test tension is "the responses of students in an assortment of testing and appraisal connections" (pp. 209). The event of test tension is affected by both characteristic demeanors and situational factors. Previously, investigate suggested that test nervousness represses test performance. According to this clear, direct model, more nervousness fundamentally deciphered into poorer execution. however, in their far reaching study on test tension, Sarason, Mandler, and Craighill [39] placed that moderate levels of uneasiness (conceptualized as the aggregate of state and characteristic nervousness) would create better test execution when contrasted with larger amounts of nervousness. Here, it is important to characterize and separate state and characteristic uneasiness:

“State anxiety has been defined as a transitory feeling of tension and apprehension; it may fluctuate over time and can vary in intensity. In contrast, trait anxiety denotes relatively stable individual differences in anxiety proneness and refers to a general tendency to respond with anxiety to perceived threats in the environment.” (Vigneau & Cormier, [50], pp. 280).

This substitute recommendation supported by Sarason, Mandler, and Craighill [39] hints at a curvilinear model of test uneasiness and test execution. This suggests the relationship between test nervousness and execution is more intricate than generally suspected. There are different elements that direct the relationship between the two. For instance, execution amid the lessons versus tests, massed versus appropriated routine of test taking, and item difficulty.

This finding can be comprehended as far as Eysenck and Calvo's [10] handling effectiveness hypothesis. On the basis of this hypothesis, nervousness may either improve or hamper undertaking execution. They trust that confinements in working memory limit are in charge of the decrements in the subjective execution of exceptionally test- anxious people. This is because in test circumstances, these people experience assignment immaterial musings, for example, stresses and dread of antagonistic results, which halfway possess working memory limit. In simple assignments, the remaining memory limit may suffice to satisfy errand necessities. In complex errands, however, it may not. Subsequently, high-anxious people will show execution decrements essentially in complex undertakings.

In this manner, well-documented evidence developed, construct fundamentally in light of correlational work, embracing this conflict of a curvilinear relationship amongst tension and performance. For instance, Rocklin and Thompson [34] found out that on a genuinely complex examination, students reporting lower test anxiety had a tendency to represent preferable scores over their high-test nervousness partners. Then again, when given a moderately simple test, understudies with unobtrusive levels of test nervousness accomplished higher scores than those keeping up either low or elevated amounts of uneasiness. Since these inquires about test the up to this point straightforward convictions with respect to uneasiness, this hypothetical viewpoint requires further examination.

III. METHODS AND MATERIAL

3.1 Participants

The population from which the subject of the present study was chosen were included 24 Arab students at Foundation English Department. They are college students who learn English as a foreign language during 3 semesters, including levels A_2, B_1, before starting they specialties. These students are in level 1 according to the university placement test.

3.2 Instrument

This study made use of interchange series (intermediate section) level A2 by Cambridge University Press and many other EFL websites for the purpose of collecting data.

The other book which is used in this procedure was ``New Headways Academic Skills``, level 1 by Oxford University Press.

3.3 Procedures

The procedures of this study cover 16 weeks of teaching English, 4 hours per day, during 5 days of a week, totally 20 hours a week. These classes include all 4 skills with same rate.Every day 3 hours for Interchange book and 1 hour for Academic Skills.

IV. RESULTS AND DISCUSSION

A Data Analysis

This part of the article examines all necessary statistics to find an answer for the research question:

4.1 Midterm Exam

Midterm exam covers 30 percent of an educational semester. Total 60 marks of this exam includes 5 sections as follow: listening (10 marks), reading (10 marks), grammar (15 marks), writing (15 marks), and study skills (10 marks).

Gathering data for midterm exam has been done through two levels: 1. Unofficial oral interview and 2. Quantitative statistics through excel and SPSS software. According to the interviews, students believed that since their teacher bombarded them with extra activities, samples of tests, and very difficult quizzes, midterm exam was so easy. Although one investigation has been done by the present researchers of this paper and it revealed that except 2 or 3 questions all remained were really easy. (Criteria were based on teachers' experience in EFL teaching)

On the basis of excel and SPSS following data has been gathered:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Midterm	24	19.9633	4.14856	.84682

One-Sample Test

	Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Midterm	23.574	23	.000	19.96333	18.2116	21.7151

As it is shown in the above table, p value is .267 which is higher than 0.05, so it means that there is no

significant difference between item difficulty and students performance.

4.2 Final Exam

Final exam covers 50% of an education semester and the arrangement of questions are as follow :

Listening (10 marks), reading (10 marks), grammar (12 marks), study skills (8 marks), writing (10 marks). Regarding oral interview with students all stated the exam as one of the most difficult one during English learning. To equalize the weight of midterm exam and final exam all scores in final changes from 50 % to 30 % and following results were gathered:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Final	24	15.1375	3.75819	.76714

One-Sample Test

	Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Final	19.732	23	.000	15.13750	13.5506	16.7244

In this table the p value is .859 which is bigger than 0.05, so there is no significant difference among variables in final exam.

4.3 Discussion

In this section some of the statistics were compared in midterm exam and the final one. According to following comparison the mean difference in midterm exam is 19.9633 while this score is 15.1375, and this shows that the total scores in final exam is lower than midterm exam.

The other option which can be discussed here is analysis of item pass and fail in midterm exam and final exam. Regarding this analysis following data has been gathered:

Midterm Exam

Section	Listening	Grammar	Reading	Study Skills	Writing
Pass	12	6	23	16	15
Fail	12	18	1	8	9
Total	24	24	24	24	24

Final Exam

Section	Listening	Grammar	Reading	Study Skills	Writing
Pass	10	14	9	11	3
Fail	14	10	15	13	21
Total	24	24	24	24	24

According to above investigations, it is revealed that test difficulty has effect but not significant on the learner's performance. The big difference is in reading (1 failure in midterm exam while 15 failures in final exam) and writing (9 failures in midterm exam while 21 in final exam). The following figures show this comparison better:

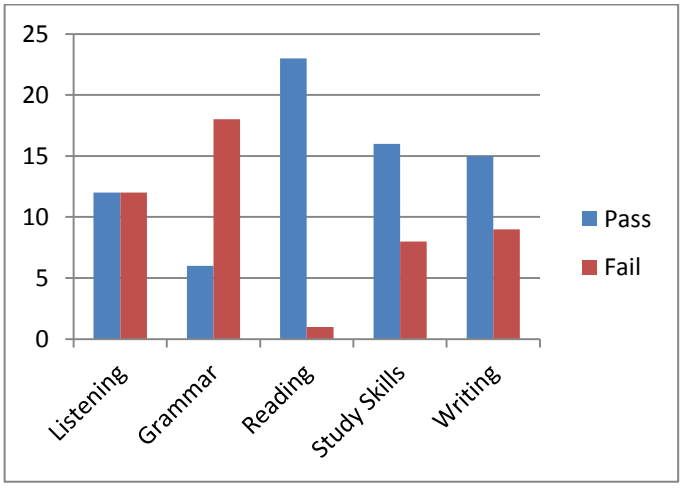

Figure 1. Midterm Exam

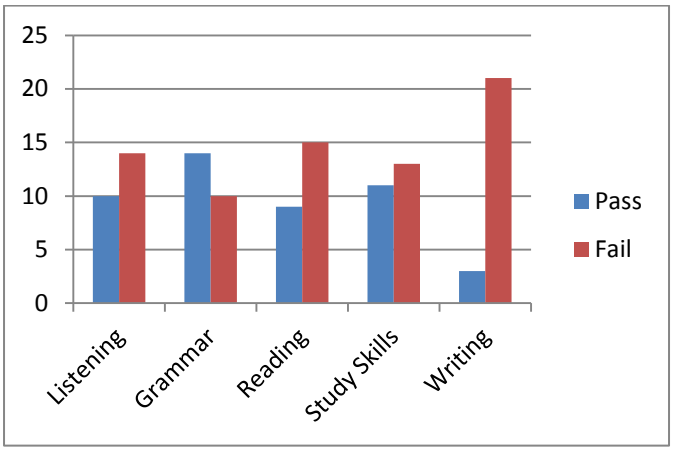

Figure 2. Final Exam

It is true that the scores are similar to somehow, but since the effect of final score is 50 percent (midterm 30 percent) thus it has effect on the learners scores.

Also Item difficulty for all items has been measured and it's a witness for truthfulness of above data:

Item Difficulty (P) = $\frac{NP}{N}$	
P Mid. List = $\frac{12}{24}$ = 0.5	**P Fin. List = $\frac{10}{24}$ = 0.41**
P Mid. Gram= $\frac{6}{24}$ = 0.25	**P Fin. Gram= $\frac{14}{24}$ = 0.53**
P Mid. Rd= $\frac{23}{24}$ = 0.95	**P Fin. Rd= $\frac{9}{24}$ = 0.375**
P Mid. S.s = $\frac{16}{24}$ = 0.6	**P Fin. Ss = $\frac{11}{24}$ = 0.458**
P Mid. Wri. = $\frac{15}{24}$ = 0.625	**P Fin. Wri. = $\frac{3}{24}$ = 0.125**

According to this table the idem difficulty differences in reading and writing are really big in final exam.

Present study aimed to measure the effect of test difficulty on learner's performance in final exam after some month training. Totally the analysis of data showed that test difficulty has no significant effect on learners score.

V. CONCLUSION

The aim of this paper was to find an answer for this question: `` does the difficulty of an exam support students poor performance in final evaluation? Finding an answer, 24 EFL students were selected and some statistical analysis were conducted on their scores in midterm and final exams of the same educational semester. The results of the study revealed that there is a slight difference between student's low performance and item difficulty but it is not significant.

Suggestions for further research

Following suggestions are made for researchers:

1. This study investigated item difficulty effect on EFL learners with small population of learners, while further research can be done with big size community.
2. For the purpose of this study EFL learners were selected, however it can be useful if in ESL contexts, such investigations happen.
3. The third suggestion and the most important one is that some researches must be done because of criteria to select question makers. I found some situations while the question maker comes from different background, he/ she is responsible to design questions in ELT context.

VI. REFERENCES

[1] Alpert, R., & Haber, R. N. (1960). Anxiety in academic achievement situations. Journal of Abnormal and SocialPsychology, 61, 207–215.

[2] Arvey, R. D., Strickland, W., Drauden, G., & Martin, C. (1990). Motivational components of test-taking. Personnel Psychology, 43, 695-716.

[3] Ball, S. (1995). Anxiety and test performance. In C. D. Spielberger& P. R. Vagg (Eds.), Test Anxiety: Theory, assessment, and treatment (pp. 107–113). Washington, DC: Taylor & Francis.

[4] Brehm, S. S., Kassin, S., & Fein, S. (2005). Social Psychology, 6 th ed. Houghton Mifflin: New York.

[5] Cale, N., Fowler, Ch., &Rempfer,M. (2012). Text Anxiety and Learning Potential in College Students. Undergraduate Research Journal for the Human Sciences. Vol. 11, 2012.

[6] Combs, H., Micheal, L., Fiore, B. & Poling, D. (2007). Easy Test or Hard Test, does it Matter? The Impact of Perceived Test DifficultyTime and Test Anxiety. Undergraduate Research Journal for the Human Sciences. Vol. 6, 2007.

[7] Eady, S. (1999). An Investigation of Possible Correlation of General Anxiety with Performance

in Eleven-plus Scores in Year 6 Primary School Pupils. Educational Psychology, 19 (3), 347–359.
[8] Ebel, R.L., &Frisbie, D.A. (1986). Essentials of educational measurement. Englewood Cliffs, NJ: Prentice-Hall.
[9] Eisenberger, R. (1992). Learned industriousness. Psychological Review, 99 (2), 248–267.
[10] Eysenck, M.W., &Calvo, M.G. (1992). Anxiety and performance: The processing efficiency theory. Cognition & Emotion, 6, 409–434.
[11] Fairclough, S. H., Tattersall, A. J., & Houston, K. (2006). Anxiety and performance in the British driving test. Transportation Research Part F, 9, 43–52.
[12] Foos, P. W. (1992). Test performance as a function of expected form and difficulty. TheJournal of Experimental Education, 60, 205–211.
[13] Foos, P.W. & Fisher, R. P. (1988). Using tests as learning opportunities. Journal of Educational Psychology, 80(2), pp. 179-183.
[14] Gurung, R. A. R. (2005). How do students really study (and does it matter). Teaching of Pschology,32(4), 239-241.
[15] Hancock, D. R. (2001). Effects of Test Anxiety and Evaluative Threat on Students' Achievement and Motivation. The Journal of Educational Research, 94, (5), 284–290.
[16] Hembree, Ray. (1988). Correlates, causes, effects, and treatment of test anxiety. Review of Educational Research, 58(1), pp. 47-77.
[17] Hong, E. (1998). "Differential stability of individual differences in state and trait test anxiety. "Learning and Individual Differences, 10,51-69.
[18] Hong, E. (1999). Test anxiety, perceived test difficulty, and test performance: temporal patterns of their effects. Learning and Individual Differences,11(4), 431-448.
[19] Hong, E., &Karstensson, L. (2002). Antecedents of state test anxiety. Contemporary Educational Psychology, 27, 348–367.
[20] Kellas, G., & Butterfield, E.C. (1971). Effect of response requirement and type of material on acquisition and retention performance in short-term memory. Journal of Experimental Psychology, 88, 50–56.
[21] Kim, S. H., & Rocklin, T. (1994). The temporal patterns of worry and emotionality and their differential effects on test performance. Anxiety, Stress and Coping: An International Journal, 7, 117–130.
[22] Kivimaki, M. (1995). Test anxiety, below-capacity performance, and poor test performance: Intrasubject approach with violin students. Personality and individual Differences, 18, 47–55.
[23] Lusk, S. L. (1981). Test anxiety, level and accuracy of predicted performance. Psychological Reports, 49, 527–532.
[24] Marso, R. N. (1969). The influence of test difficulty upon study efforts and achievement. American Educational Research Journal, 6 (4), 621–632.
[25] Matlock-Hetzel, S. (1997). Basic Concepts in Item and Test Analysis. Annual Meeting of the Southwest Educational Research Association, Austin, January, 1997.
[26] Meijer, J. (1996). Learning potential, personality characteristics, and test performance. In Hamers, J. H. M., Sijtsma, K. &Ruijssenaars, A. J. J. M. (Eds.) Learning Potential Assessment: Theoretical, Methodological, and Practical Issues. (pp. 341 – 360). Steenwijk, Netherlands: GrafischProduktiebedrijfGorter.
[27] Meijer, Joost. (2001). Learning potential and anxious tendency: test anxiety as a bias factor in educational testing. Anxiety, Stress, and Coping, 14, pp. 337-362.
[28] Miksch, K. L. (2003). Legal issues in developmental education: the impact of high-stakes testing. Research and Teaching in Developmental Education, 19 (2), 53–58.
[29] Nami, Y. (2006). The effects of test anxiety on listening test performance. System, 34, 317–340.
[30] Pintrich, P. R., Smith, D. A. F., Garcia, T., &McKeachie, W. J. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). Educational and Psychological Measurement, 53, 801-813.
[31] Putwain, D. (2007). Researching academic stress and anxiety in students: some methodological considerations. British Educational Research Journal, 33 (2), 207–219.
[32] Qayyum, R. & Ismail, Z. (2011). The Effects of Immediate Forewarning of Test Difficulty on Test Performance in Applied Educational Context. International Journal of Business and Social Science. Vol. 2, No.1, 2011.
[33] Rempfer, M., Hamera, E., Brown, C., & Bothwell, R. (2006). Learning proficiency on the

Wisconsin Card Sorting Test in people with serious mental illness; What are the cognitive characteristics of good learners? Schizophrenia Research, 87, 316-322.

[34] Rocklin, T., & Thompson, J. M. (1985). Interactive effects of test anxiety, test difficulty, and feedback. Journal of Educational Psychology, 77, 368–372.

[35] Rocklin, T. (1997). Self-adapted testing: Improving performance by modifying tests instead of examinees, Anxiety, Stress & Coping, 10 (1), 83–104.

[36] Sapp, M. (1993). Test anxiety: Applied research, assessment, and treatment intervention. Lanham, MD: University Press of America.

[37] Sapp, M. (1996). Three treatments for reducing the worry and emotionality components of test anxiety with undergraduate and graduate college students: Cognitive-behavioral hypnosis, relaxation therapy, and supportive counseling. Journal of College Student Development, 37, 79-87.

[38] Sarason, I. G. (1961). Test anxiety and the intellectual performance of college students. Journal of Educational Psychology, 52(4), pp. 201-206.

[39] Sarason, S. B., Mandler, G., &Craighill, P. G. (1952). The effect of differential instructions on anxiety and learning. Journal of Abnormal and Social Psychology, 47, 561–565.

[40] Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: Reactions to tests. Journal of Personality and Social Psychology, 46, 929–938.

[41] Sax, G., Eilenberg, E. G., & Klockars, A. J. (1972). Achievement as a function of test item complexity and difficulty. The Journal of Experimental Education, 40, 90–93.

[42] Sax, G., & Reade, M. (1964). Achievement as a function of test difficulty level. American Educational Research Journal, 1b (1), 22–25.

[43] Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. Anxiety Research, 4, 27-41.

[44] Spielberger, C. D. (1980). Preliminary professional manual for the Test Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press.

[45] Spielberger, C. D. (1983). Manual for the State-Trait Anxiety Inventory Form V. Palo Alto, CA: Consulting Psychologists Press.

[46] Spielberger, C. D. (2009). Test Anxiety Inventory. InI. B. Weiner & W. E. Craighead (Eds.), The Corsini Encyclopedia of Psychology. John Wiley & Sons, Inc.

[47] Thiede, K. W. (1996). The relative importance of anticipated test format and anticipated test difficulty on performance. The Quarterly Journal of Experimental Psychology, 49A (4), 901–918.

[48] Tobias, S., & Everson, H. T. (1997). Studying the relationship between affective and metacognitive variables. Anxiety, Stress, and Coping, 10, 59–81.

[49] Trifoni, A. &Shahini, M. (2011). How Does Exam Anxiety Affect the Performance of University Students?Mediterranean Journal of Social Sciences. Vol.2, No.2. 93-100.

[50] Vigneau, F., & Cormier, S. (2008). The factor structure of the state-trait anxiety inventory: An Alternative View. Journal of Personality Assessment, 90 (3), 280–285.

[51] Waugh, N. C. (1967). Presentation time and free recall. Journal of Experimental Psychology, 73, 39–44.

[52] Wiedl, K. H. (1999). Cognitive modifiability as a measure of readiness for rehabilitation. Psychiatric Services, 50(11), pp. 1411-1419.

[53] Weber, C. J., & Bizer, G. Y. (2006). The effects of immediate forewarning of test difficulty on test performance. The Journal of General Psychology, 133 (3), 277–285.

[54] Zajonc, R. B. (1965). Social facilitation. Science, 149, 269-274.

[55] Zeidner, M. (1998). Test anxiety: The state of the art. New York: Plenum Press.

[56] Zimmer, J. W., & Hocevar, D. J. (1994). Effects of massed versus distributed practice of test taking on achievement and test anxiety. Psychological Reports, 74, 915–919.