WCES-2011

The Relationship between Test Anxiety, Motivation and MI and the TOEFL iBT Reading, Listening and Writing Scores

Mohammad Amiryousefi a *, Mansoor Tavakoli b

a ACECR Institute of Higher Education (Isfahan Branch), Isfahan, Iran
b Department of English, Faculty of Foreign Languages, University of Isfahan, Isfahan, Iran

Abstract

TOEFL iBT is widely used as a certificating device, and is strongly claimed accountable by the people who utilize it. However, there are indications of the vulnerability of TOEFL iBT to the test takers’ characteristics. The purpose of the present study is to provide empirical evidence to see if there is a relation between test anxiety, type of motivation and intellectual abilities of test takers and their scores on TOEFL iBT reading, listening and writing sections, specially the verbatim use of the input texts in the integrated writing task which is read and listen to write. To carry out the present study a Multiple Intelligences Development Assessment Scales (MIDAS) questionnaire, a test anxiety scale and some survey questions detecting test anxiety provoking factors were given to 30 subjects attending TOEFL iBT classes at ACECR. Then their scores on a TOEFL iBT test were compared, and the language features in their written texts were analyzed for source use to find out the possible effects of the given characteristics. The results revealed that first, test takers usually had test anxiety second, some factors like time limitation and lack of self-confidence provoked test anxiety and third, there was a relation between musical and kinesthetic intelligences and writing and listening respectively. Linguistic analysis of their responses to the integrated writing task, however, revealed no instances of the verbatim use of the source.

© 2011 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.

Keywords: Test anxiety, test anxiety provoking factors, MI, verbatim use of the input texts

I. Introduction:

Language testing is mainly concerned with the extent to which the obtained scores truly mirror test takers’ underlying abilities in a particular area in a specific testing situation. It, therefore, deems necessary to answer some questions before any attempts are made to devise a test. These questions are: What is it that we are testing? What tasks and items can elicit it? In what format should the tasks and items be presented to the test takers? In what mix should the items be included to represent the ability or the construct appropriately? What scoring procedure should be taken to be able to extrapolate the test performance to the underlying behavior? (Weir, 2005; Fulcher & Davidson, 2007). Throughout the past decade, test developers and experts have devoted much of their interest and energy to develop a theoretical view of language ability to investigate the nature of the language proficiency, and to develop and apply more sophisticated statistical tools to analyze language tests and test takers’ performance to tap these issues more profoundly (Bachman, 1991).

* Mohammad Amiryousefi. Tel.: +98.091.321.323.81; fax: +98.031.139.125.41
E-mail address: M.Amiryousefi@Yahoo.com.
However, research in language testing has indicated that language ability is not the only thing that affects test takers’ performance on a language test. Personal traits like motivation, anxiety, ambiguity tolerance etc. that test takers bring to the testing situation can affect their performance too (Bachman, 1990, Skehan, 1989, 1998; Tavakoli, 2009). These construct-irrelevant factors, as Messick(1996) puts it, are considered the potential sources of test bias which can make the obtained scores unrepresentative of the underlying ability that a language test wants to measure, and put the whole testing process at stake (Takala & Kaftandjieva, 2000; Messick, 1996). A thorough examination of the probable effects of such factors is, therefore, worth investigating.

Taking into account these considerations and the status of TOEFL iBT as a proficiency test around the globe, the purpose of this study is to measure the probable impacts of test anxiety, multiple intelligences and motivation on listening, reading and writing scores of this test.

2. Theoretical background

Test fairness is an aspect of language testing which is hotly debated in the literature. The central aim of test fairness discussions is to make tests free from bias and to contribute to testing equality (Kunnan, 2007). A test may be considered biased when students having the same language ability perform differently. Different factors like content or language of the test, when the content of the test is biased to test takers from certain groups, religions, ethnicities etc.; standards of the test that may affect test takers from different groups differently; and test takers’ characteristics such as age, learning strategies and styles, attitude and motivation, aptitude and intelligence and anxiety can contribute to test bias (Kunnan, 2007, 2004; Weir, 2005; Brown, 2004). These factors can affect the validity of a test and cause measurement errors. A major concern in the design and development of language tests is, therefore, to minimize the effects of these factors that are not part of the language ability (Bachman and Palmer, 1996; Messick, 1989, 1996).

2. 1. Test anxiety

Test anxiety is a form of anxiety concerning apprehension over academic evaluation which comes from a fear of failure (Horwitz and Young, 1991). It is believed to be caused by certain factors like time limit, test techniques, test format, length, testing environment and clarity of test instructions (Young, 1999). Some researchers believe that test anxiety negatively affects language learning process and the test outcomes. It is believed to cause cognitive interference while preparing for examination, taking examination or both. Test anxiety during the preparation for a test may lead to poor understanding and organization of the concepts and hence make the retrieval cumbersome. While test anxiety during the test decreases test takers’ attention and thereby increases the number of errors (Ohata, 2005; Cassady & Johnson, 2002; Sarason, 1980, 1986). It seems that test anxiety inversely affects test performance. Chapell et al. (2005), however, believe that for some students some anxiety might be a good thing. It might motivate longer periods of study and more careful attention to questions on the exam.

2.2. Multiple Intelligences

One of the most compelling, yet controversial new approaches to education reform is Multiple Intelligences Theory, or MI. Recent advances in cognitive science, developmental psychology and neuroscience suggest that each person’s level of intelligence is actually made up of autonomous faculties that can work individually or in concert with other faculties. Howard Gardner originally identified seven such faculties, which he labelled as “intelligences”: Musical Intelligence; Bodily-Kinesthetic Intelligence; Logical-Mathematical Intelligence; Linguistic Intelligence; Spatial Intelligence; Interpersonal Intelligence and Intrapersonal Intelligence. Recently, he added an eighth intelligence to the list: the Naturalist Intelligence (Visser et al., 2006).

Loori (2005) believes that preferences of intelligences are at work in educational settings. The results of his study indicated that males are more successful in tasks that involve logical and mathematical intelligences, whereas females are more successful in tasks involving intrapersonal intelligence. Intelligence type can, therefore, affect learning and testing. Mahdavy (2008), for example, believes that linguistic intelligence significantly contributes to the listening proficiency and Richards and Rodgers (2001) believe that musical intelligence contributes to English pronunciation. Razmjoo (2008), however, found no significant relationship between language success and the types of intelligences in particular.
2.3. Motivation

Motivation is defined as a derive, a desire or an emotion which fuels L2 learners to strive to learn (Gardner, 1985; Brown, 1994). L2 motivation has three major components motivational intensity; a desire to learn; and a positive attitude (Purpora, 2004). According to Gardner, there are two types of motivation: integrative and instrumental. An integratively motivated L2 learner shows interest in learning about the culture and the people of the target language. Integrativeness is a desire to learn a language in order to ‘come closer to the other language community’ (Gardner, 2001: 5). Integratively motivated learners are believed to invest more effort on the learning process and enjoy it more (Lamb, 2004). An instrumentally motivated learner, on the other hand, has more pragmatic considerations in his/her mind regarding L2 learning, such as obtaining a job or earning more money. An instrumentally motivated learner approaches language learning for utilitarian objectives (Tavakoli, 2009).

Motivation is studied as a factor interacting with test takers’ performance and affecting test results. Researchers are now concerned with finding what kind of motivation leads to purposeful engagement of test takers’ abilities on a language test. Some researchers, however, claim that integrative motivation is more advantageous (Lamb, 2004; Kunan, 1995).

2.3. The purpose of the study

Taking the effects of factors such as anxiety, motivation and MI on test performance and the status of TOEFL iBT around the world, the purpose of this study is to seek the answers to the following questions:

Q1: Is there a relation between test anxiety, MI and motivation and test takers’ performance on TOEFL iBT listening, reading and writing tests?
Q2: Do gender and marital status moderate this relation, if there is any?
Q3: Is there any relation between these factors and the verbatim use of the source in the writing of the test takers in the integrated writing task of TOEFL iBT?

3. The study

3.1. Participants

Thirty subjects (8 male and 22 female) attending TOEFL iBT classes at ACECR participated in the study. Their age ranged from 20 to 36. They had learned English for at least two years but had not received any TOEFL iBT preparation courses. They were tested at the beginning of the TOEFL iBT preparation course to neutralize the effects of coaching. They were all at the intermediate level according to the results obtained from both the OPT (2004) and TOEFL iBT (2006).

3.2. Procedure and Instrumentation

First Oxford Placement Test (2004) was given to 45 participants and 30 subjects who had the same level of proficiency were chosen. A week later a TOEFL iBT test released by ETS in 2006 was administered. Then a Multiple Intelligences Development Assessment Scales (MIDAS) questionnaire (in Likert scale type ranging from always=1 to never=0); a Test Anxiety Scale adapted from Sarason’s (1984) (in Likert type always=5, usually=4, sometimes=3, rarely=2, never=1) containing 22 multiple-choice items and some survey questions detecting test anxiety provoking factors; and Gardner's Motivation Test Battery (MTB) (Gardner, 1985) (adapted to a 5-point scale, ranging from ‘Strongly Agree’ to ‘Strongly Disagree’) were given to the subjects. As the integrated writing task in TOEFL iBT test is read and listen to write-i.e. students read a short passage, listen to a short lecture and then are asked to respond to a writing question in about 200 words, the procedures followed by Cumming et al. (2006) and Ohkubo (2009) were finally used to analyze their responses to see if there is any relation between test anxiety, MI and motivation and the verbatim use of the source (use of the input text or lecture without reformulation or reference to the source). The performance of the test takers were scored on the basis of the scoring procedures used by ETS for TOEFL iBT test. For the writing section, however, three experienced raters were recruited to rate the writing tasks. They were all experienced English teachers and were aware of TOEFL scoring procedures. To obtain inter-reliability, an average of the three raters’ scores was used to arrive at the final score.
4. Results

Table 1 represents the descriptive statistics of the test takers’ scores. To ensure sample homogeneity Kolmogorov-Smirnov Test was used. As shown in Table 2, $a = 0.05 > \text{Sig} = 0.562$ which represents sample homogeneity. Then stepwise regression was used to explore the relationship between the scores and the given factors. As indicated in Table 3, $\text{Sig} = .027$ is smaller than $a = 0.05$ in Writing $= \alpha + (\beta \times \text{Musical})$ model representing the relation between musical intelligence and writing, and $\text{Sig} = .007$ is smaller than $a = 0.05$ in listening $= \alpha + (\beta \times \text{Kinesthetic})$ model representing the relation between kinesthetic intelligence and listening. For other factors no relation was found. The results also indicated that factors such as gender and marital status brought about no difference.

Table 1. Descriptive Statistics of the Scores obtained

<table>
<thead>
<tr>
<th></th>
<th>writing</th>
<th>reading</th>
<th>listening</th>
<th>total mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>N Valid</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Missing</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mean</td>
<td>1.57</td>
<td>1.53</td>
<td>1.37</td>
<td>2.9828</td>
</tr>
<tr>
<td>Std. Error of Mean</td>
<td>.092</td>
<td>.093</td>
<td>.089</td>
<td>.30314</td>
</tr>
<tr>
<td>Median</td>
<td>2.00</td>
<td>2.00</td>
<td>1.00</td>
<td>.144</td>
</tr>
<tr>
<td>Std. Deviation</td>
<td>.504</td>
<td>.507</td>
<td>.490</td>
<td>.504</td>
</tr>
<tr>
<td>Variance</td>
<td>.254</td>
<td>.257</td>
<td>.240</td>
<td>.254</td>
</tr>
<tr>
<td>Skewness</td>
<td>-.283</td>
<td>-.141</td>
<td>.583</td>
<td>-2.062</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-.206</td>
<td>-.127</td>
<td>-.1784</td>
<td>-2.127</td>
</tr>
<tr>
<td>Sum</td>
<td>47</td>
<td>46</td>
<td>41</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Kolmogorov-Smirnov Test

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Normal Parameters(a,b) Mean</th>
<th>Std. Deviation</th>
<th>Most Extreme Differences Absolute</th>
<th>Asymp. Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>writing</td>
<td>30</td>
<td>2.9828</td>
<td>.30314</td>
<td>.144</td>
<td>.562</td>
</tr>
<tr>
<td>reading</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>listening</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. The Results Obtained from Stepwise Regression

<table>
<thead>
<tr>
<th>Model</th>
<th>Dependent Variable</th>
<th>R</th>
<th>R Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Regression writing</td>
<td>.404(a)</td>
<td>.163</td>
<td>5.465</td>
<td>.027(a)</td>
</tr>
<tr>
<td>2</td>
<td>Regression reading</td>
<td>.373(c)</td>
<td>.139</td>
<td>.307</td>
<td>.970(c)</td>
</tr>
<tr>
<td>3</td>
<td>Regression listening</td>
<td>.484(b)</td>
<td>.234</td>
<td>8.550</td>
<td>.007(b)</td>
</tr>
</tbody>
</table>

a. Predictors: (Constant), musical; b. Predictors: (Constant), kinesthetic; c. Predictors: (Constant)

The analysis of the responses of the test takers to the integrated writing task, however, revealed no instances of verbatim use of the source. But those subjects who showed higher level of test anxiety tended to use the source mostly by making references to the text or lecture by the use of reporting verbs and expressions like “as stated by the writer”, “according to the lecturer”, “the writer believes”.

Finally, the questions provided to detect the sources of test anxiety showed that time limit, length of the test and lack of confidence were amongst the most prevailing causes of test anxiety.

5. Conclusion and Discussions

This study was an effort to empirically explore the effects of test anxiety, MI and motivation on the performance of the test takers on listening, reading and writing sections of TOEFL iBT test. The findings of the study revealed that there is a relation between musical and kinesthetic intelligences and writing and listening respectively. People who have musical intelligence are able to hear patterns, recognize them and manipulate them. They can become good writers and composers. People with kinesthetic intelligence, on the other hand, mostly involve motor rather than cognitive abilities and seem to use muscle memory (they remember things through their body). These people can become good musicians, actors, doctors etc. (Gardner, 1983; Wikipedia, 2010). It can be concluded that TOEFL iBT test might be biased toward test takers with a base of musical and kinesthetic intelligences.

The linguistic analysis of the writing of the test takers in the integrated writing task, however, revealed no instances of the use of the ideas in the input text or lecture without reformulation or clear references. The answers of
the test takers to the questions exploring the possible causes of test anxiety ranked time limit, length of the test and lack of confidence among the most common sources of test anxiety. Time management strategies can, therefore, be helpful in this regard.

References