ELIMINATE MEASUREMENT ERROR ATTRIBUTABLE TO TEST ANXIETY

Huang Tao  Zhuang Xuhong
Xichang College

Abstract
This paper illustrates students' test anxiety attributable to test itself from four aspects of testing procedure—planning, preparing, administering and scoring. It is the test anxiety that impedes effective and successful testing performance of students and further causes measurement error. All educators endeavor to eliminate measurement error and expect no biased results or decisions based on students' observed scores. Therefore suggestions to solve such problems are provided here.

Key words
eliminate measurement error test anxiety testing procedure

Any test has an error of measurement, the discrepancy between students' true scores and their observed scores. On occasion of gross measurement error, a number of educational decisions based on their observed scores are somewhat biased. Neither teachers nor students expect the biased results or decisions. However, the causes of measurement error are complex. One of the main reasons is the test anxiety which students experience when taking a test.

For some of them, tests are burdensome and they experience an abnormal sense marked by physiological and behavioral accompaniments. This abnormal feeling or emotional state is called test anxiety. When students experience test anxiety, worry and emotionality are called into play to impede effective and successful testing performance. Dusek 1989. Then test scores can not reflect accurately their proficiency of learning or what they have learned. Both of students and teachers hope that the negative effect of test anxiety would be mediated in that students' true scores can be gained fairly.

The phenomenon of test anxiety was first studied in the 1920s. Since then, a diverse set of hypotheses, theories, and principles has been developed to describe and analyze this phenomenon. As reported in a number of research studies, social factors, personal factors, instructional factors, as well as the test itself lump together to facilitate the test anxiety. The purpose of this paper is to focus on examining how the test itself puts test anxiety into effect so that teachers, especially those teaching English as a foreign language, can gain a general understanding of it and also modify some aspects of the test to help students affected by test anxiety.

The first aspect to affect the test anxiety within the test itself is associated with planning procedure.

Before examining this procedure, I like to draw your attention to the distinction between a classroom test and a quiz. A classroom test is announced in advance and covers a specific unit of instruction. It may be given constantly every two or three weeks—in some cases every week and require a considerably longer time than a quiz. The quiz may be unannounced to check what has been covered in the previous lesson or period. It may be given even every period at the beginning or end of the class and cover only a few minutes. Usually a quiz includes easier questions than a test.
Bearing in mind this distinction, let’s first look at the effect of the test announcement on test anxiety. When a teacher enters a classroom and administers a test which is not announced in advance, high-anxious students may respond immediately in their mind by raising questions like this: “Is the test difficult?” How can I do if I do not prepare adequately?” This kind of worry usually results in “attentional blocks,” extreme concern with autonomic and emotional sensations, and cognitive deficits such as misinterpretation of information” Dusek 1989 [6] that cause poor performance. This is why a test is supposed to be announced in advance to allow students to have an adequate preparation. On the other hand, one may ask why a quiz may not be announced. The answer may lie in the finding of Caron research. He proved that there was a significant difference for an easy test between high-anxious and low-anxious students. Phillips et al. 1989 [32]. Thus, we are safe to say that an unannounced quiz may not contribute a lot to arousing student test anxiety since they know it is easy and only covers a very short time.

Meanwhile, with frequent unannounced quizzes, students are encouraged to devote time regularly to their studying and reviewing. The value lies in its positive effect on students’ learning habits and affords opportunity to practice the art of test taking which helps mediate test anxiety often affecting performance on longer tests. Valette 1977 [21]; Yue Xiaodong. An interesting study of test anxiety by comparison between Chinese and American high school students proved that American students were more anxious than Chinese students. The lower test anxiety of Chinese students was attributed to the practice effect due to their overexposure to tests and quizzes in preparation for the College Entrance Examination.

The consistent tests and quizzes, however, should not be administrated randomly at teachers’ willingness. We call for a long-term plan to use these consistent periodic quizzes and tests in an aimed, coherent, and systematic way. In an aimed way, they should reflect directly the focal course contents and objectives covered in a certain period. In a coherent way, they should be related to each other and the previous knowledge would better be slightly embodied in the successive tests. In a systematic way, they should as a whole reflect what has been taught with weight during the long period and the students’ knowledge can be integrated through practice. Thus, through the careful plan of the periodic quizzes and tests, students are frequently aware of the instructional objectives and the emphasis in the course as well as their own weakness and those language items needing remedy. Learning is fostered by frequent review and remedial work. The adequate study integrates students’ knowledge over a long period and helps students build their confidence for the final examination or more important tests.

The second aspect to affect the test anxiety within the test itself is associated with preparing procedure, which usually involves writing test items and assembling the test paper.

When preparing test items, some amateur teachers think it is too simple to choose one from among overwhelming language testing sources or just too easy to take pen in hand and turn out items without much consideration. Except by chance, such an item will not prove a very valid measure and usually enhance students’ anxiety. Consider the following two test items applied to elementary students:

1] On seeing the young child fall into the lake, John jumped _______ the water immediately and went to his rescue. Multiple Choice Item
   A] at  B] into  C] on  D] for

2] The first step is often the most difficult. Substitution Item

In item 1, Heaton 1988 [97] structure “On doing sth’” and the word “rescue” are more difficult than the language point tested — “jump into.” Students have to focus more attention to the understanding of the stem’s situation provided by the sentence than the language point itself tested. In item 2, all four choices are much more difficult than the word “first” tested. This encourages the students to close their minds to all else but the choice to be guessed. Harris 1969 [56]. If a test includes quite a number of such items with poor quality, it will not benefit the student’s motivation and confidence involved in the test and further arouse the worry about their uncertain choices even if they know the language points tested.

Another point deserving our attention is the test bias that can easily arouse some students’ test anxiety. For example, in reading comprehension, certain words have far greater currency among others.
The words — debit, account, balance, margin, erase, file — are biased towards students learning business. Culture-loaded words may also cause some troubles such as Laundromat, sunny side up, three credit hours which may be unknown to some students. Thus the same guessing problems affecting their scores occur to those who are not in that specific field or with cultural disadvantage. Therefore when a test item functions differently against or for a particular group of students, it may be considered biased in that factors associated with students or a test that are not related to the trait being measured. Such test bias can be minimized through sensitivity review to identify offensive language or biased content. Unless a test is purposely devised for a specialized group, the test writer should endeavor to restrict the content problem in a general way.

Fortunately a number of guidance books on testing English as a foreign language have been published through decades. They afford language teachers principles and techniques to develop a language test. So the first step to avoid writing bad items is to equip yourself with such principles and techniques. However, testing is a practical issue. Before you become a competent teacher in testing, you have to practice a lot to gain abundant experience.

Notice that the above two test items as suggested are only suitable for students at elementary level but not at intermediate level if they are well revised. The item difficulty therefore becomes another crucial important variance in testing consideration. Rocklin and Thompson found that a hard test will benefit the low anxious students for they did relatively better on a hard test than on an easier one. The moderately anxious students did the opposite and the most highly anxious students did poorly on both hard and easy test. That is to say, the moderately anxious students are likely to be less anxious on an easier test than on a hard one because they may have more confidence on an easier test. Then when we choose or write the items we must bear in mind the question — “Is this item too difficult or too easy for my students?” And generally speaking, the difficulty level of the whole test should be based on the average proficiency of all test takers so that it may benefit as many students as possible. The trick here is how to decide the difficulty level of an item. Most teachers usually do this by experience and some statistic-sophisticated teachers use statistical techniques to achieve this in a scientific way.

After choosing or writing of the test items, we should assemble the test papers. But in what order should we arrange the items in the final test paper? Lund found that students got better scores on tests with items sequenced by increasing difficulty than tests with some difficult items put early in the test. It is apparent that the easy items put early in the test may provide a rewarding experience that students show how well they can handle specific elements of the language learned and help create positive attitudes toward the successive items which are more difficult. Conversely, some difficult items put early in the test decrease the students confidence as they move on. And anxiety may be further aroused to interfere with the easier items put in the later part. However, Kestenbaum and Weiner did not find significant correlation between item sequencing and test anxiety in their research. In my opinion, this may be due to some students good test taking techniques. When they meet the problems they cannot deal with, they just skip it so as to concentrate on the items they can do. In short, item sequencing of the test has sometimes been found to make a difference and it is safe therefore to arrange the test items with increasing order of difficulty.

The third aspect to affect the test anxiety within the test itself is associated with administering procedure.

When we administer the test, any issues related to the mechanics of testing may enhance inadvertently the student anxiety level. Consider for instance a quiet location in which students are taking a listening comprehension test as compared to another group taking the same test with the distraction of noise. If all other factors are administered constant, which group do you think will be less anxious and do better? Likewise, if the quality of the equipment to administer the same listening comprehension test to different groups is not the same, the difference is made here to arouse some students’ anxiety. Therefore, a responsible teacher should always check all the possible mechanics of testing before administering the test.

When we administer the test, timing the test becomes another potential problem. Flass and Hill
found that low anxious students performance was facilitated under time pressure situation while high anxious students performance was facilitated under no time pressure condition. 1989 [40] Apparently the removal of time pressure resulted in strong optimizing effects for anxious students. In reality however it is almost impossible to have no time limits. It is obvious that students test anxiety will be aroused if time is given too short. On the other hand it may seriously impair the validity of the test to distinguish among students if time is given too long. Then how long should the test be administrated so that it can benefit as many students as possible. Any competent teacher knows this question is related to the item content, item number and item difficulty included in the test paper. We must always make balance between the test contents, objectives, item number, item difficulty and time limits. That is to say the content with much weight should have a large proportion of item numbers and rating and therefore much time is to be provided for them. And the items with difficulty should be allowed more time.

What more when considering the time provided for the test we should reserve the time for administrating procedures such as seating the students and handing out the materials giving general directions collecting the materials at the end of the testing period handling unanticipated problems and include time for students to read the directions and adjust mentally to each new section accompanied. In most common classroom testing situations it would probably be fairest and safest to time the tests so that all but the slowest 10 or 15 percent of the students are able to attempt all the items. (Harris 1969 [119])

The fourth aspect to affect the test anxiety within the test itself is associated with scoring procedure. In a face to face interview rater attitude toward the students exerts an important effect on their test anxiety. For instance after the student provides certain initial replies the rater cues the student with questions like “Can you think of other reasons?” “Is there any more do you want to add?” Thus the student usually provides another answer to the item and it is likely to increase the students score. Other forms of cueing like compliments also serve to “instruct” students. Phillips et al. 1989 [3] We do not know however whether raters tend to prompt anxious students more than other students. But it is true that some teachers tend to cue students whom they like better than others. Thus it is fairest to suggest that teachers other than the student teachers are assigned to rate the face to face interview.

In an objective test there exists test anxiety attributable to scoring schemes which correct for guessing. A study by Sherriffs and Boomer showed that if a test was announced to have a “penalty” for guessing students who lacked confidence in their own judgment were easily threatened by the announcement of “penalty” and omitted more items. It thus resulted in test anxiety and even more doubt about the answers they were sure at one time. Under this situation anxious students performed more poorly. Phillips et al. 1989 [32] Therefore it is better not to adopt the “penalty” instructions for correction of guessing but adopt other statistical techniques in order to achieve a better test validity.

The foregoing illustrations summarize the research studies done by other educational researchers as well as some of my own comments on test anxiety attributable to test itself. The affecting aspects exist in every testing procedure — planning, preparing, administrating and scoring. Teachers should have a clear mind about all these possible aspects and try to minimize them while proceeding to test students so that it can help students affected by test anxiety and also gain fair and true scores of the students that are relatively free of measurement error caused by test anxiety.

References


Continued on p81