Validation of objective-type test in biology at secondary school level

Syed Imtiaz Hussain Naqvi\textsuperscript{a},* , Muhammad Aamir Hashmi\textsuperscript{b}, Akhtar Hussain\textsuperscript{a}

\textsuperscript{a}Ph. D Scholar University of Education, Lahore, 54000 Pakistan
\textsuperscript{b}Lecturer, IER, Punjab University, Lahore, 54000 Pakistan

Received November 1, 2009; revised December 9, 2009; accepted January 18, 2010

Abstract

Evaluation should not only depict the real state of students’ achievements but also help to improve teaching learning process. Test should be more valid and reliable. The system of the examination and the nature of the test in Pakistan need a thorough change and improvement.

In order to improve the situation the Federal Ministry of Education developed objective type tests in the subject of Biology but these tests have not been tried out in the field to establish uniformity of standard, reliability and validity, although need for such a practice was felt at the time of development and publication of these tests. In the present study, a part of these has been tried out for the purpose of education.

A number of 48 students of two schools in Lahore were taken as sample. Scores were tabulated and item difficulty and item discrimination index indices were calculated. The results of the study show that the test on the whole has exhibited poor validity as far as the difficulty level and discriminating index are concerned. It has been found out that the tests need revision and must be tried out on a larger sample.

Keywords: Validation; objective type tests; biology; secondary school.

1. Introduction

In the present system of education in Pakistan, the main purpose of examination is assigning of grades and classification of students. On the basis of these grades, tests are designed to find those students who fail at the end of a particular course, those who succeed and those who get second or first division. This is a very narrow and restricted purpose of evaluation. Education, today, is seen as a process in terms of behaviour attitudes and skills. Evaluation as a component of the education system must contribute to this end and must help to improve the teaching - learning process. The method of evaluation should consist of acquiring and processing the evidence needed in determining the students’ level of learning and effectiveness of the teaching. It should be an aid in clarifying the significant goals and objectives of education and process for determining the extent to which students are developing along these desired ways. It must also provide feedback about the effectiveness of the teaching-
learning process. The system of evaluation and examination has been under discussion by the people concerned i.e. educators and there has been a consensus on its change. This doesn’t provide any reasonable judgment about students’ ability and development of skills.

1.1. Review of Related Literature

Education, conceived as a system, has many components and different aspects. It may be influenced by a variety of factors and function in different ways. However, from whatever angle it is looked upon, the product of this system ‘is in the form of some change in the behaviour of the students. At the end of a certain period of instruction we are interested in knowing the extent to which the change has occurred. We also need to describe and appraise the change... Thus evaluation in education may be seen as:

“The systematic collection of evidence to determine whether intact certain changes are taking place in the learners as well as to determine the amount or degree of change in individual students”. (Bloom 1981)

1.2. Evaluation in Education

Evaluation means the process of “putting value on” or ‘assigning worth’ to some thing (Clark 1981). It tells us about the student achievement; at the end of a particular teaching-learning process, estimates of the quantity and quality of pupils learning and the positive or negative effects of the pertinent factors influencing the teaching learning process.

It also provides us with the information about the merits and de-merits of our pupils and educational programmes in the light of our instructional objectives.

Process of evaluation is completed in two steps. In the first step data is gathered for an assessment of pupils, status. In the second step the information obtained through measurement (assessment) is used to make measurable judgment of the achievement level. Thus the essential element in evaluation is judgment and so it is more inclusive and comprehensive term than measurement, It includes quantitative and qualitative description (i.e. measurement or assessment) plus a value judgment.

Measurement, a part of evaluation, provides ‘information’ upon which an evaluation can be used. It is the quantitative description of students attainment. Describing merely the situation, is measurement but judging its value, is called evaluation.

1.3. Purposes of Evaluation and Measurement

Evaluation has many purposes are teachers use it as a basis for school marks, reporting to parents and promotion, administrators use it as a basis for categorizing pupils in to groups, guidance counselors use it as a basis for pupils advice, pupils use it as basis for mapping out their own programmes, educational experts use it as a basis for curriculum revision, but its most important role is its use in the teaching learning itself because:

It gives the teachers the feed—back they need in order to know what the pupils have learnt and what to do next. It gives the pupils the feed—back they need in order to profit from their successes

1.4. Types of Tests

On the basis of abilities that are to be tested in the learners, the tests can be categorized as follows: Aptitude test, Attitude tests, Personality tests, Intelligence test, Achievement tests, Standardized tests, Teacher made tests. Ebel (1979) Thorndike Isa Khan

1.5. Comparison of standardized test and teacher-made tests:

Standardized tests are based on general features, content and objectives common to many school through out the country. Teacher-made tests are built for particular requirements of an individual class in a particular school
Standardized tests deal with large segments of knowledge or skill. Teacher made tests are often designed for a much smaller area of knowledge or skill, or even for a single topic. Standardized tests are constructed and developed by experts with enough resources at their disposal. Teacher-made tests depend on the limited skill and slight resources of one or two teachers. Standardized tests provides norms for various groups often on a natural scale. Teacher-made tests provides norms only on a specific school scale.

Essay test, Objective test, Multiple choice items, True false items matching items, Completion items  Harry Schofield (1972)

1.6. Characteristics of a good test

A test may be called a good one if it has properties are Clarity, Reliability, Practicability, Objectivity, Validity, Clark (1981)

1.7. General rules for objective test construction:

The item should clearly pose a single definite problem. The reading and linguistic difficulty of items should be appropriate for the examinees. Each item should be brief and avoid repetition. Simple words should be used as far as possible. The grammar and punctuation of the items must be faultless. Clues to the correct responses should be avoided.

1.8. Steps for improving tests:

Steps are generally while improving a test are Try out, Assessment of content validity, Grouping on the basis of scores, Formulation of items statistics i.e. Difficulty level, Discrimination index Bloom (1981).

1.9. Procedure

The test was administered in Govt Iqbal Hussain High School, Ghari Shahoo, Lahore and Govt N. D. Islamia High School, Ichhra, Lahore. A random sample of 24 students of 9th class from each of the above mentioned school was taken. The tests were marked and the achievement scores of each student were tabulated in a descending order. The item analysis was done for the computation of difficulty level and discrimination index. Difficulty level for each test item was calculated by the formula.

\[
\text{Difficulty Level} = \frac{H + L}{N} \times 100
\]

Where H – stands for higher achiever 
L – stands for lower achiever 
N – stands for total number of students.

### Tables

<table>
<thead>
<tr>
<th>RANGE</th>
<th>REMARKS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 – 19%</td>
<td>Difficult</td>
</tr>
<tr>
<td>20 – 49%</td>
<td>Less Difficult</td>
</tr>
<tr>
<td>50 – 60%</td>
<td>Average</td>
</tr>
<tr>
<td>61 – 80%</td>
<td>Easy</td>
</tr>
<tr>
<td>81 – 100%</td>
<td>Easiest</td>
</tr>
</tbody>
</table>

Discrimination index for each test item was calculated by the formula.

\[
\text{Discrimination Index} = \frac{H - L}{N/24}
\]
2. Analysis of Data & Results

Test items developed by Ministry at Education Islamabad for Secondary classes were applied to 4-8 students of 9th class of the following two schools.

Test covered the chapter on “classification of living organism and comprised of 133 items including all 5 types of objective type tests.
Test was marked and test items were analysed and tabulated for computation of difficulty level and discrimination index.

3. Difficulty Index

<table>
<thead>
<tr>
<th>Range</th>
<th>Remarks</th>
<th>Test items number</th>
</tr>
</thead>
<tbody>
<tr>
<td>0% to 19%</td>
<td>Difficult</td>
<td>Nil</td>
</tr>
<tr>
<td>20% to 49%</td>
<td>Less difficult</td>
<td>8,9,12,26,3,35,37,57,72,74,75,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>78,80,82,90,92,98,110,123,125,126,131,132</td>
</tr>
<tr>
<td>50% to 60%</td>
<td>Average</td>
<td>2,7,13,14,17,22,45,50,54,59,61,70,71,73,77,79,95,96,101,103</td>
</tr>
<tr>
<td>60% to 80%</td>
<td>Easy</td>
<td>1,5,6,10,11,15,16,18,19,20,21,23,24,25,27,28,30,34,36,38,39,40,41,43,44,46,47,48,49,51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>,52,53,55,56,58,60,61,62,65,66,67,69,76,81,83,85,87,88,89,93,94,99,100,102,107,109,1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13,114,115,117,118,124,133.</td>
</tr>
<tr>
<td>81% to 100</td>
<td>Easiest</td>
<td>3,4,29,32,33,42,63,68,84,86,91,97,104,105,106,111</td>
</tr>
</tbody>
</table>

4. Finding

The analysis of the data revealed that:

1. Zero percent of the test item fall in the category of difficulty level ranges from 0% to 0.19%
2. The difficulty of the 17 percent items number ranges from 20% to 49%.
3. The difficulty of 22 percent of the test items number, ranges from 50% to 60%.
4. The difficulty of 47% of the test item number, falls between 60% to 80%.
5. The difficulty of 2 percent of items numbers, ranges between 81% to 100.
6. Discrimination index of zero percent of the test items in ranges from 0.81 to 1.00.
7. 7 percent of the test items number 6,15,27,41,69,113,126,128,130,132 have discriminating index between 0.51 to 0.80.
8. 50 percent of the test items number have discriminating index between 0.21 to 0.50.
9. 29 percent of the test items number have not discriminating index between 0.00 to 0.20.
10. 12 percent of the test index number 1,8,14,16,59,76,89,95,97,98,99,106,110,114,123,125,131 shows negative values.
5. Conclusions

On the basis of the findings of the study, it was concluded that:

1. About 1/8 test items of the total items are either very difficult or very easy.
2. Most of the items are satisfactory with regard to the difficulty level.
3. About 1/3 of the total items are such that do not discriminates between the examinees.
4. About 1/8 test items are ambiguous.
5. Half of the items are either more or less discriminating.
6. Considering difficulty level and discriminating index only, the test on the whole exhibit poor validity.

6. Recommendation

It is recommended that:

1. The test items which are very easy should be excluded
2. Nothing any test items number being very difficult who may be revised.
3. The test items which do not discriminate between the examinees. They may be excluded.
4. The test items which seem to be ambiguous should be revised.

7. Suggestions

It is suggested that:

1. In the light of findings and recommendation of this study the test should be revised and re-tested.
2. The test should be completely validated with respect to table of specifications, determination of co-relation and experts views.
3. The test may be tried out on a larger sample of students population.

References