Construction and Standardization of Three Tier Concept Achievement Test CAT in Science

Dr. Manmeet (Baewja) Oberoi
Professor cum Principal
Shah Satnam Ji College of Education, Sirsa, India

Abstract:
The paper presents the details of the development and standardization of the Concept Achievement Test (CAT) in Science to measure the errors and misconceptions among secondary school students. The test was constructed and standardized by the researcher including the concepts of: “adaptations, habitat, biosphere, ecosystem, food chain and food web, functions of ecosystem, biomass and biodiversity”, from the unit of environment in PSEB prescribed science text book of ninth class. The preliminary draft contained 53 items and was administered to two hundred and twenty students from the government schools of Punjab. On the basis of the indices of item difficulty (0.16-0.25) and in view of judges 23 items were discarded and 30 items were selected for the final draft. The final draft of three tier test was administered to one hundred and twenty five students. Since this test was constructed to differentiate errors from misconceptions so scoring of the three tier test was unique. If the student gave right response in first tier, right reason in second tier and confident in tier 3, total score for the question was 1. If in any one of the 3 tiers the student got 0 marks, the score was 0. In this way score of each student for all the three tiers was calculated. The reliability of the test was found to be 0.79 and was calculated by Kuder Richardson reliability coefficient. Validity was established by content validity method.

Keywords: Concept Achievement Test (CAT) in Science, Errors, Misconceptions, Three Tier Test

I. INTRODUCTION

Achievement test can be defined as a test designed to measure the effects of specific teaching or training in an area of the curriculum. OR It is a standardized test used to measure acquired learning. Anderson (1972) was of the view that: "By 'achievement test' I mean a set of questions asked to test what a person has learned from exposure to instruction." In case of psychological testing, the psychometric properties to be taken care of while constructing and standardizing an achievement test determine its efficacy and applicability to measure learning outcomes of students at a particular stage or grade during the process of learning. Gronlund (1985) is of considered view that "when we focus our measurement of learning tasks, criterion-referenced measurement is quite feasible and is preferred because of its descriptive nature." To study the errors and misconceptions among secondary school students the Concept Achievement Test (CAT) in science in the content area of environment was constructed and standardized by the researcher to include the concepts of: “adaptations, habitat, biosphere, ecosystem, food chain and food web, functions of ecosystem, biomass and biodiversity”, on the basis of PSEB prescribed science text book of ninth class. The construction and standardization of the Concept Achievement Test (CAT) in science was based on the methodologies devised and used by researchers to study misconceptions and errors in different subjects/specific area of content matter (Adeniyi, 1985; Trowbridge and Mintzes, 1988; Amir and Tamir, 1990; Schmidt et al., 1991; Odom et al., 1995; Griffard and Wandersee, 2001; Ozam, 2001; Eryilmaz and Surmeli, 2002; Cataloglu, 2002; Ozay and Oster, 2003; Sindhu and Sharma, 2004; Cetin et. al., 2004; Demircioglu et. al., 2005; Haki, 2005; Kutluay, 2005).

The content was categorized under seven headings - Adaptations, Habitat, Biosphere, Ecosystem, Food Chain and Food Web, Functions of Ecosystem, Biomass and Biodiversity. In order to develop test, the content area was divided into smallest units and expressed in behavioural objectives in terms of taxonomy of cognitive domain (Bloom, 1968), keeping in view the six major classes of cognitive domain i.e. Knowledge, Comprehension, Application, Analysis, Synthesis and Evaluation.

Selection of Appropriate Test Item Format
Since the purpose of the Concept Achievement Test (CAT) in science was to identify errors and misconceptions, a three tier test was constructed as explained below. The first tier of each item in the test was in a multiple choice format asking the content knowledge of the students with four choices (Rollnick and Mahooana, 1999). The students are to give the reason in second tier for the answer given in the first tier (Treagust, 1988; Odom and Barrow, 1995; Tan, Goh, Chia and Treagust, 2002). In the third tier students are asked whether they are confident about the answer or not. This format discriminated the students’ lack of knowledge from the misconceptions by means of third tier items which assessed how confident the students were about their responses for first and second tier (Eryilmaz and Surmeli, 2002; Cataloglu, 2002; and Kutluay, 2005).

Scoring of the Concept Achievement Test (CAT) in science
Scoring was done by recording the student's answers on the raw data as incorrect answer or correct answer.

Tier I: If the answer is correct in first tier, the student will get one mark; if the answer is incorrect, the student will get 0 marks.
Tier-2: If student gives right reason for the answer in first tier, he/she will get 1 mark in tier 2 otherwise he/she will get 0 marks.

Tier-3: If the student is confident about the answer given will get 1 mark otherwise 0.

If the student gives right response in first tier, right reason in second tier and confident in tier 3, total score for the question will be 1. If in any one of the 3 tiers the student gets 0 marks, the scores will be 0. In this way score of each student for all the three tiers were to be calculated.

**Preparation of Preliminary Draft of the Test**

After the planning stage, the next step was to prepare the preliminary draft of the test. To facilitate this procedure and establish the content validity of the test, a blueprint of the test, in terms of two way chart called "table of specifications" was drawn in the beginning. The table of specifications related the instructional objectives to the course content and specifies the relative emphasis to be given to each type of learning outcome. Gronlund (1985) has given three steps for preparing a table of specifications: obtaining a list of instructional objectives; outlining the course content; and preparing the two way chart. Following this the test items were prepared on the basis of the table of specifications in consultation with the prescribed textbook, the question bank, the teachers teaching the concerned subject and the experts in the field of test construction, mainly from the faculty of education and psychology. While preparing the test items for the preliminary draft of the test, it was ensured that these are in accordance with the objectives and are capable of measuring as accurately and objectively as possible the behaviour reflected by objectives. Every precaution was taken while framing the items so as to make them clear, ambiguous, direct and precise. A set of general as well as specific instructions for each type of item was also prepared to facilitate accurate responding on the part of the students. The preliminary draft of the test along with the set of instructions was shown to the language experts in order to eliminate ambiguity in the language of the items, if any.

**Content Validation**

The items were shown to the judges (the teachers teaching science subject in schools and language experts) and were asked about the comprehensiveness of items, representativeness of the content by the items, and appropriateness of the format. The subject experts investigated the test items in terms of whether the items were appropriate for the 9th class students, the misconceptions intended to assess and whether the items were easily understood.

**Evaluation of the Test**

No matter, how painstakingly the test items are prepared for the preliminary draft, there is no guarantee that they will operate in the same manner as planned. It was therefore, desirable to evaluate the test items on various criteria in order to find out how they will work with the group of students for whom they have been prepared and to judge the quality of the test in general. For this following strategy was adopted.

**Difficulty Value:** The difficulty value of each item was computed by adopting the procedure of finding the proportion of students who had given correct answer for the item and determining the item difficulty by computing the formula

\[
\text{Difficulty value (D.V.)} = P \times q
\]

Where \( P \) stands for proportion of students in the total group scoring the item correctly; and \( q \) stands for proportion of students in the total group scoring the item incorrectly. For finding the difficulty value of the items, the test was administered to 220 students from Govt. schools of Patiala affiliated to Punjab school Education Board.

**Selection of Items and Preparation for Final Draft**

On the basis of the indices of item difficulty (0.16-0.25) and in view of judges 23 items were discarded and 30 items were selected.

<table>
<thead>
<tr>
<th>No</th>
<th>Draft</th>
<th>Objectives in % age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Knowledge</td>
</tr>
<tr>
<td>1</td>
<td>Preliminary</td>
<td>26.5</td>
</tr>
<tr>
<td>2</td>
<td>Final</td>
<td>26.6</td>
</tr>
</tbody>
</table>

**Preparation of Final Draft**

Final draft of Concept Achievement Test (CAT) in science was prepared with 30 items having three tiers. The first tier had multiple choice questions; the second tier having the reason for choosing the answer in the first tier; and the third tier showing the confidence level of the students while responding in first two tiers of Concept Achievement Test (CAT) in science.

**Administration and Scoring of Final Draft**

Now this test was administered to 125 students of 9th class in various secondary schools of Patiala District to find out the reliability and validity of the test. Every precaution was taken to make the test administration effective. Time limit for taking the test was 1 hr and 20 minutes; so that all the students can attempt all the items. The test booklets were scored according to the scoring key and the obtained data was tabulated for further analysis. Scoring of the Concept Achievement Test (CAT) in science was done by recording the student’s answers on the test as incorrect answer or correct answer’ as discussed in preliminary draft. If the student gave right response in first tier, right response in second tier and confident in tier 3, total score for the question was 1. If in any one of the 3 tiers the student got 0 marks, the score was 0. In this way score of each student for all the three tiers was calculated and also for first tier, first two tiers and all the three tiers will be calculated.

**Reliability and Validity of Concept Achievement Test (CAT) in Science**

Reliability is the consistency of the scores obtained. A test is considered meaningful if it produces consistent results. The
reliability of the test was calculated by calculating the coefficient alpha which is a measure of internal consistency of an exam. This coefficient is a general form of Kuder Richardson reliability coefficient (Fraenkel and Wallen, 1996). The reliability of the test came out to be 0.79, which was considered to be satisfactory measure of the reliability of the Concept Achievement Test (CAT) in science. Validity refers to the appropriateness, meaningfulness and usefulness of the specific inferences researchers make based on the best results (Fraenkel and Wallen, 1996). Content validity of the Concept Achievement Test (CAT) in science was established. Content validity refers to the content and format of the test. The purpose of the content validation is to assess whether the items adequately represent a performance domain or construct of specific interest i.e. how well the test portrays the domain of concepts it is intended to represent.

II. IMPLICATIONS
The methodology adopted in the construction and standardization of Concept Achievement Test (CAT) in Science will provide a lead to future researchers, not only in the area of science, but in other subject areas as well. Therefore, it is expected that the present paper will open new vistas for researchers to provide new direction in teaching-learning process along with providing some concrete suggestions for educational practitioners to adopt such strategies in school education. Moreover, the psychological principles underlying three tier testing in concept achievement will provide a new orientation to evaluation of learning outcomes of students in a more comprehensive manner. Assessing students’ misconceptions is very important for effective learning.

III. REFERENCES