Development and Standardization of Achievement Test in Accountancy

Ekta Mehta
Assistant Professor Malwa Central College of Education for Women, Ludhiana, India

Abstract

The paper reports details of the development and standardization of the test of achievement in accountancy meant for class XII under PSEB. The test initially consists of 78 items. Later after discussion with experts, the items were reduced to 57. After thorough item analysis the final draft of the test reduced to 35 items. The test-retest reliability was found to be 0.76. Validity was established by content validity method.

INTRODUCTION

An information and technology based society requires individuals who are able to think critically about complex issues, analyze and adapt to new situations, solve problems of various kinds and communicate their thinking effectively. The study of bookkeeping and accountancy equips students with knowledge, skills and habits of mind that are essential for successful and rewarding participation in such a society. To learn accounting in a way that will serve them well throughout their lives, students need classroom experiences that help them develop accounting understanding; learn important facts, skills and procedures; develop the ability to apply the accounting processes and acquire a positive attitude towards accountancy. The misconception about achievement in any subject is that it is student’s skill in remembering the content area of the subject that they study. Tests in achievement are therefore prepared to assess student’s knowledge in the subject matter. The concept of achievement in accountancy has a dual goal of achieving the subject matter competency and developing cognitive skills. Achievement signifies accomplishment or proficiency of performance in given skill or body of knowledge.

The achievement tests are controlling process for teaching learning activities. They help in evaluating the effectiveness of teaching instructions. According to Dowine (1974) any test that measures the attainments or accomplishments of an individual after a period of learning is called achievement test. It can also be thought of as a sample of indicator of a student’s knowledge taken at a particular point in a time.

Therefore achievement tests propose to measure what and how much pupils have learnt as a result of formal and informal instructions. These are utilized as evaluating courses of study or efficiency of teacher and teaching methods or other educational factors. Lindeman (1967) classified achievement tests in following categories:

- Teacher Made Test
- Standardized Test

Teacher made tests are frequently the basis of evaluating the students progress in school. These are more specifically focused and they usually reflect the content of a particular unit or course. The teacher made test is tailored to measure the achievement of students and intended objectives for them after completing a series of learning tasks for the course.
On the other hand standardized tests are particular kinds of test, different from the final examination a high school teacher might design for his course. When talking about tests, "standardized" simply means that everyone who takes the test is given the same amount of time and sees the same or very similar test questions. "Standardized" also means scoring is done very carefully so that the test scores do not depend upon who happens to be doing the scoring. These tests are widely used because, by and large, they have shown to be an efficient way to collect information about what people know and can do. These are objective in the sense that test scores do not depend to any great extent on who happens to score the answers. These are valid in that they often provide relevant and useful data for making decisions about mastery of a body of knowledge and skills and potential for success.

The need was felt to develop an achievement test in accountancy for assessing the achievement of the students in the subject at the senior secondary stage i.e. XI level as there was lack of suitable achievement test in accountancy among the existing tests.

CONSTRUCTIONAL METHODOLOGY OF ACHIEVEMENT TEST

Achievement test in accountancy was developed with the objective to evaluate the knowledge, understanding and application of the content matter achieved by senior secondary class students. Before constructing the achievement test, text material pertaining to XI class under PSEB system was thoroughly analyzed. Brainstorming sessions with various commerce teachers; with minimum five years’ experience was conducted; to seek their views to plan requirements of construction.

Purpose of constructing the test

The purpose of test was to measure the acquisition of accounting concepts. The test was made to measure Knowledge, Understanding and Application of accounting concepts by commerce students at senior secondary level.

Target Population

This achievement test is meant for the target population of XII class students who had just passed XI class exams under Punjab School Education Board, successfully.

Type of Test Items

All the items were of multiple choices. Each item had four alternatives out of which one is correct. After studying thoroughly the syllabus prescribed for XI class accounts (PSEB), the main content area was identified and the major course instructional behavioural objectives were specified. Adequate provisions were made for evaluating all the important outcomes of instruction. Then an initial draft consisting of 78 items was prepared. While writing the items following points were kept in mind:

- Items were made strictly according to accountancy syllabi prescribed by PSEB for XI class.
- The items were phrased in such a manner that there was no ambiguity regarding their meaning.
- The items were not concerned with the trivial aspect of the subject matter, that is, it only measured the significant aspects of knowledge and understanding.
- Interdependence among the items was avoided.
- The number of items selected was three times greater than to be retained in the final draft.
The keyed responses were randomly varied in position. The items written were reviewed, criticized and revised after thorough discussion with experts for the content validity. Hence following the suggestions of experts, 21 items were deleted and a few of them were modified. The distracters of multiple choice type items were distributed equally so that they should also be in a position to discriminate between students. The preliminary achievement test consisted of 57 items, a few easy and a few difficult. All items of a particular content were placed together.

For the present test a brief summary of concepts with total number of items derived from them was prepared as under:

**Table I: Concept-wise distribution of items in the Preliminary Draft**

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Name of the Topic</th>
<th>No. of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Accounting Concepts</td>
<td>11</td>
</tr>
<tr>
<td>2.</td>
<td>Journal</td>
<td>10</td>
</tr>
<tr>
<td>3.</td>
<td>Ledger</td>
<td>9</td>
</tr>
<tr>
<td>4.</td>
<td>Trial Balance</td>
<td>8</td>
</tr>
<tr>
<td>5.</td>
<td>Final Accounts</td>
<td>10</td>
</tr>
<tr>
<td>6.</td>
<td>Bank Reconciliation Statement</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>57</strong></td>
</tr>
</tbody>
</table>

**Scoring**

A scheme of evaluation was prepared for scoring the papers. One point of credit was given for each correct response.

**Try Out of the Test**

The test was administered to a sample of 25 students. During data collection proper rapport was established and maintained with the students. The respondents were explained how to take the test. Every reasonable precaution was taken to ensure normal conditions during the administration of the test.
After preliminary try out, problems faced by the students were noted and given the due consideration at the time of revision of first draft. On the basis of performance of the students out of 57 items, 10 items were clearly discarded and in three items slight changes were made to improve clarity in the wording. Finally 47 items were retained which formed the first draft of the test.

**Final try-out of the test**

The success of any test depends upon the efficiency with which items of the test were constructed, which is reflected in the values of difficulty and discriminative power of the item. For item analysis, the test was administered on a sample of 100 students of class XII (commerce) of PSEB. Then after working out difficulty value and discriminative power of every item, 35 items were selected in the final draft.

Objective wise distribution of various items in the final draft is given in Table II

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Topic</th>
<th>Weightage</th>
<th>Total No. of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Knowledge</td>
<td>Understanding</td>
</tr>
<tr>
<td>1</td>
<td>Accounting Concepts</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Journal</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>Ledger</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>Trial Balance</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>Final Accounts</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6.</td>
<td>Bank Reconciliation Statement</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>12</strong></td>
<td><strong>13</strong></td>
</tr>
</tbody>
</table>

**Item Analysis**

Item analysis is a technique through which those items which are valid and suited to the purpose are selected and the rest are either eliminated or modified to suit the purpose. The main objectives of item analysis are enumerated below:

- It indicates which items are difficult, easy, moderately difficult or moderately easy. In other words, it provides an index of the difficulty value of each item.
• It also provides indices of the ability of the item to discriminate between high and low. In other words, item analysis indicates the discrimination value of each item. This is known as discrimination index.

For the present test 27 percent upper and 27 percent lower cases were considered for calculating difficulty value and discrimination value of each item. Kelley (1939) showed that the product-moment correlation between a test item score and the total score could be estimated by using only the tails of the distribution and he also showed that the most efficient division to use was the top and bottom 27 percent tails. This suggestion was followed. The valued test papers were arranged from the highest score to the lowest score. The top 27 percent of the students were kept in one group. This was the high group. The poorest 27 percent of the students were kept on the other group. This was the low group. The middle group consisting of 46 percent of the papers was kept aside since the two extreme groups high and low only were needed for item analysis.

For calculating the difficulty value (D.V.) and discriminating power (D.P.) of items following formulas were used:

\[ D.V. = \frac{R_U + R_L}{N} \]
\[ D.P. = \frac{R_U - R_L}{N/2} \]

Where

\[ R_U = \text{Total Number of Right Responses in the Upper Group} \]
\[ R_L = \text{Total Number of Right Responses in the Lower Group} \]
\[ N = \text{Total Number of Students in Both the Groups} \]

**Final Draft of the Test**

According to Garrett (2007) “As a general rule, validity indices of .20 or more are regarded as satisfactory”. Following this suggestion the items with difficulty value from .25 to .75 and discriminating power from .22 to .75 were retained in the final draft of the achievement test. On the basis of this a total of 35 items were selected in the final draft.

**Reliability**

Reliability refers to the consistency of scores or measurement which is reflected in the reproducibility of the scores. In the words of Anastasi (1951), Reliability refers to the consistency of scores obtained by the same individuals when re-examined with the same test on different occasions or with different sets of equivalent items or under other variable examining conditions.

Mouly (1970) remarked, “The test-retest method is the only feasible approach to the establishment of the reliability of the test. The answer to one question given by a respondent in two instances can be compared for estimating consistency.” Keeping this in view, the test-retest method was found better option as assumptions of unifactor test and parallel items were not met and the split half method was discarded because items were not arranged on the basis of item type.

The test was administered on a sample of 35 students who were not included in the experimental sample of population. With a gap of fifteen days the test was
again administered to the same group. The coefficient of reliability was found to be 0.76.

Validity

Validity can be best defined as the degree to which a test is capable of achieving certain aims. The validity of the test is determined by measuring the extent to which it matches with a given criterion.

Cronbach (1960) says that validity may be determined by showing “That a test corresponds to the definition of the trait intended to be measured, or it may be established inductively by naming the traits represented in the items in hand.”

Regarding establishing the validity of a test Mouly (1970) remarked, “At the most elementary level, it is necessary for all the tests to have content validity i.e. each question must be related to its topic. There must be an adequate coverage of the overall topic. The questions must be clear and unambiguous.”

In the present study the questionnaire responses of the test were validated against the competencies. Content validity was found by relating each item carefully against the competencies. To confirm validity the test items along with a list of competencies to be developed in the learners were given to a panel consisting of five experts in the subject matter and two in test items. Comments on item clarity, its correspondence to outcomes along with validity and objectivity of data were used to make modifications in some items. Only those items were kept in test draft for which there was complete agreement among the experts. The content validity was ascertained by preparing a blue print of the test items indicating the weightage given for the objectives and competencies by these items.

References