Title:

DEVELOPMENT OF CRITERION-REFERENCED TEST IN ENGLISH FOR STANDARD IX

The problem:

Traditional tests measure achievements in reference to one's class, group, school or state for passing judgement. This type of norm-referencing is useful to gather information about an examinee in relation to his relative ability or relative attainment but it lacks information regarding an examinee's ability or inability in a specific area. It is also insufficient to say about the individual performance of an examinee. So the present situation of measurement has led us to the realisation of the need for the development and use of the Criterion-Referenced Tests (CRT).

A CRT assesses an individual's performance level with respect to a well-defined domain of behaviours. It emphasises test organisation based on specific tasks or behavioural objectives and assessment in terms of predefined performance criteria.

Development of CRTs is at initial stage in different areas of testing. English Grammar is one of them and so the present study was conducted specially for the development of CRTs in English (Grammar) for Standard IX (Middle Level).
The purposes of the study were:

(i) to develop CBIE in English (Grammar) for Std. IX,
(ii) to determine the cut-off score of the test,
(iii) to establish the reliability of the tests,
(iv) to validate the tests, and
(v) to locate the examinees into mastery - non-mastery stages.

Procedure:

For the development of the tests, twelve steps suggested by Hambleton were adopted. They were:

1. Preliminary considerations,
2. Domain specifications,
3. Item-writing,
4. Assessment of content validity,
5. Revision to test items,
6. Field Test administration,
7. Revision to test items,
8. Test assembly,
9. Selection of standard (cut-off score),
10. Pilot test administration,
11. Preparation of manuals, and
12. Reliability and validity.

First of all, purposes of the study were specified and a group to be measured was selected. For the judgemental review work, qualified staff was also selected.
The content area was specified following the syllabus of the Class IX and the total content was grouped into fourteen domains. Then domainwise learning outcomes (specific objectives) were formulated adopting papham's description of 'Application'. Learning outcomes were reviewed by qualified staff appointed for the purpose.

For learning outcomes based on grammatical concepts, rules and principles, two multiple items with three responses were generated adopting the 'Theory Based Method'. The typical rules for multiple choice items suggested by Conole and O'Neil were followed.

Items for the tests were either selected or revised or replaced by more appropriate ones utilizing 'judgmental Review procedure' suggested by Humbleton, and two parallel forms 'A' and 'B' were prepared.

The tests were administered to the group of forty students to be measured before and after the instruction. Domain-wise scores of each time was found out.

Using the obtained score, the optimal cut-off score of the tests was determined. The Reliability of Domain Score and 'Reliability of Mastery Classification' were established using 'kappa index'. For the validation of the tests, the 'Content Validity' and the 'Criterion-related Validity' were established. Finally, the test manual was prepared.
Findings:

- Optimal cut-off scores:

'Empirical Method' was followed and 'Optimal Cut-off Score' was established for each of the tests. The following table shows the optimal cut-off scores of the fourteen tests.

<table>
<thead>
<tr>
<th>No. of the Domain</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Items</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td>12</td>
<td>10</td>
<td>16</td>
<td>20</td>
<td>10</td>
<td>15</td>
<td>10</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Cut-off Form-A</td>
<td>6</td>
<td>12</td>
<td>17</td>
<td>10</td>
<td>8</td>
<td>15</td>
<td>17</td>
<td>8</td>
<td>11</td>
<td>12</td>
<td>8</td>
<td>12</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>Score Form-B</td>
<td>8</td>
<td>12</td>
<td>17</td>
<td>10</td>
<td>8</td>
<td>15</td>
<td>17</td>
<td>8</td>
<td>11</td>
<td>12</td>
<td>8</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
</tbody>
</table>

- Reliability:

Hambleton's formula was applied to establish 'Estimated Domain Score' of the tests. The 'Estimated Domain Score' lay between 0.00 and 1.00 and the standard error of measurement ranged from 0.00 to 0.11, 0.00 to 0.12, 0.00 to 0.13, 0.00 to 0.14, and 0.00 to 0.16 when domains contained 20, 18, 15, 12, and 10 items.

For the reliability of Mastery Classification Decision, the category suggested by Swaminathan was selected and 'Kappa Index' was adopted. The obtained value of 'K' of both the forms lay between 0.45 and 0.89. Thus, all fourteen tests were reliable.

- Validity:

The content validity was established by using the 'Judgmental procedure'. Only good, very good and excellent