5.0.0 Introduction

In the foregoing chapters, pre-requisites of the process of standardisation of a test have been described in detail. After the establishment of norms, the test constructor should present the evidence of reliability and validity which are the two essential qualities of a good instrument. Therefore, it is obligatory on the part of the investigator to establish the reliability and the validity of the test.

5.1.0 The Concept of Reliability

The prime purpose of the measurement is to arrive at some standard and precise judgement about the individual. The judgement would be of some value, if it is based on dependable scores earned on a dependable test. The dependable test means a reliable test. The term reliability denotes trustworthiness or consistency. Ebel has given an operational definition of the test reliability as follows:

"The reliability coefficient for a set of scores from a group of examinees is the coefficient of correlation between that set of scores and another set of scores on an
This definition implies that reliability is not a property of a test by itself but rather of a test when it is applied to a particular group of respondents. The more appropriate the test is to the level of abilities in the group, the higher is the reliability of the scores it yields.

This operational definition specifies that high correlation coefficient is a good measure of reliability. One of the properties of the correlation coefficient is that it provides a relative, rather than an absolute measure of agreement between the pairs of scores obtained from the same persons. If the differences between scores for the same person are relatively small to the differences between scores for different persons, the test tends to show a high reliability. But if the differences between scores for the same person are relatively large to the differences between persons, the scores will show low reliability.

The operational definition calls for two independent measures obtained from equivalent tests of the same trait.

for each member of the group. The above discussion reflects only one aspect of reliability, namely adequacy of statement sampling. Another aspect of reliability is concerned with temporal stability i.e. the extent to which the original scores could be reproduced on a different occasions. Remmers and Gage have defined reliability as:

"The consistancy with which a test yields the same results in measuring whatever it does measure."\(^2\)

The definition indicates that there should be stability of the scores for the same individual, if the test is given repeatedly. In other words, the reliable test gives approximately the same result on two different occasions. This does not mean that there would not be any difference on two successive scores of the same individual obtained at two different occasions. There would be slight fluctuations. These fluctuations would not affect the reliability of the test.

This aspect of reliability indicates that the test should be stable with respect to the factors which operate during the interval time between the test and the re-test to the extent that the two sets of scores are correlated to that extent the test is reliable. If this

correlation is low or insignificant it means that the items are affected by the factors operating in the time interval. However, full proof reliability is still a question since the test is to deal with human beings having variety of moods at different time interval. This does not mean that one can be flexible with the 'reliability' aspect of the test.

Robert Lado points out:

"If the scores of the students are stable, the test is reliable, if the scores tend to fluctuate for no apparent reason the test is unreliable." 3

Anastasi in a simple explanation:

"Reliability prefers to the consistency of scores obtained by the same individuals when re-examined with the same test on different occasions." 4

In short a test must show sufficient evidences of its reliability. A test without the statement of its reliability would be of little value. Anastasi has rightly said:


"In spite of optimum testing conditions however, no test is a perfectly reliable instrument. Hence, every test should be accompanied by a statement of its reliability."\(^5\)

5.2.0 Methods of Estimating Reliability

There are different methods of establishing the reliability of the scale. They are:

1. Equivalent-Form
2. Test-Retest
3. Split-Half
4. Rulon
5. Flanagan
6. Analysis of Variance

5.2.1 Equivalent-Form Method

This procedure is very simple. From the very beginning the investigator has to prepare two equal forms of the test. These two forms must be very close in similarity.

They should be close in the matter of content, trait to be measured, processes required for responding the statements and the number of statements. The statements must have equal discriminative power, internal consistency and

\(^5\) Ibid., p. 77.
unidimensional ability. The examinee takes one form of the test and then the other form soon after that. In order to control some error variance or practice effect, the turns of the form should be related. The agreement between the two is determined by means of a correlation coefficient. This method overcomes the limitations of effect of time interval between two successive administrations of the evaluation device. This method is rarely used since it is very difficult to have two parallel forms of the same test. Construction of the equivalent forms needs a lot of time on the part of the investigator. Therefore this method has not been adopted for the purpose of determining the reliability of the test.

5.2.2 Test-Retest Method

The simplest obvious method of obtaining repeated measures for the same individuals of the same ability is to give the same test twice. The agreement between the two sets of scores from the two applications of the same tests is determined by means of a correlation coefficient. There are certain factors like practice, confidence, growth, physical facilities which play a certain role on those two different occasions of test administrations. To counteract the effects of these variables, a fairly large sample would be needed. Besides, the time interval should be adequate. The investigator after studying the pros and
cons of this method, used it for estimating the reliability of the present test. The scores obtained on two different administrations of the same test were used as two sets of scores for finding out the correlation coefficient. The correlation co-efficient was computed by the product moment method. The table 5.1 presents the distribution of the two sets of scores.

Table 5.1
THE SCATTER DIAGRAM OF SCORES OBTAINED BY THE STUDENTS ON CTA ON TWO SUCCESSIVE ADMINISTRATIONS OF THE TEST

<table>
<thead>
<tr>
<th>First Administration</th>
<th>0-49</th>
<th>50-</th>
<th>100-</th>
<th>150-</th>
<th>200-</th>
<th>250-</th>
<th>300-</th>
<th>350-</th>
<th>399</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>350-399</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>300-349</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>250-299</td>
<td>5</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>200-249</td>
<td>14</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>150-199</td>
<td>22</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>100-149</td>
<td>20</td>
<td>36</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>50-99</td>
<td>15</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>0-49</td>
<td>35</td>
<td>59</td>
<td>41</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td>148</td>
</tr>
</tbody>
</table>

Total 35 59 41 9 3 1 148

r = .87 ± .05

It is observed from the table 5.1 that the correlation co-efficient between the two sets of scores is +.87
and \( SE_r \) is .05. In this method the selection of the most appropriate time is important. If the time between the two occasions is too long, some significant changes taking place during the time interval would produce low correlation. Similarly if the time interval is very short there would be memory effect on the second administration. That is why the interval should not too long or too short. If this is not done, the resulting effect would produce a spurious correlation.

5.2.3 Split-Half Method

The practical difficulties associated with the determination of reliability coefficient by the test-retest and equivalent forms could be overcome by split-half method. One of the practical alternatives is to split up the entire form of the test into two reasonably equivalent halves. For making two equivalent halves, usually pooling the odd numbered statements for one set of scores and pooling the even numbered statements for the other sets of scores is done. This usually constitutes the two sets of scores obtained from a single testing and it reasonably controls such factors as practice, fatigue, distraction and mental set. After the test has been given to a representative sample of subjects, two scores are obtained for each respondee, one on the odd numbered and the other on the even numbered statements. The agreement between these
scores on the same test as determined by a correlation co-efficient, reflects the reliability of half the test. The reliability of the whole test could be obtained by Spearman Brown's Formula which is as under:

\[ r_{tt} = \frac{2 r_{1/2}}{1 + r_{1/2}} \]

The correlation between the two sets of scores was computed using the product moment method. The table 5.2 presents the scatter diagram and reliability co-efficient of the half test.

Table 5.2

<table>
<thead>
<tr>
<th>Scores on the even numbered items</th>
<th>0- 20- 40- 60- 80- 100- 120- 140- 160- Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-19</td>
<td>32 57 41 8 6 2 1 148</td>
</tr>
</tbody>
</table>

\[ r = .84 \pm .06 \]
After establishing the reliability of the half test the reliability of the whole test was obtained by using Spearman Brown Prophecy. The reliability obtained for the half scale was .84 whereas the reliability for the whole test was .91. The reliability of the whole test is quite high and it provides indication that the test is quite reliable and dependable.

5.2.4 Rulon Formula

Rulon has developed a simple formula for establishment of reliability of the total test scores using the basic definition of reliability.

"Reliability is the proportion of true variance in a test". 6

Rulon equation expresses the complementary statement that reliability is equal to unity minus the proportion of error variance. Hence computing error variance is a main thing in his work. The formula is given below: 7

\[ r_{rr} = 1 - \frac{\sigma_d^2}{\sigma^2} \]


The investigator has applied this method for finding out the reliability of the present scale. The scores, obtained by 148 pupils on two halves of the scale were used for this procedure. The difference between two halves, the square of the difference between two halves, total of the two halves and square of the two halves, of each subject is given in table 5.3.

Table 5.3
DIFFERENCE OF ODD AND EVEN NUMBERED ITEMS ALONG WITH THE TOTAL SCORE, MEAN AND S.D. FOR ESTIMATION OF RELIABILITY WITH RULON AND FLANAGAN FORMULA

<table>
<thead>
<tr>
<th>Candidates</th>
<th>Score of odd numbered items</th>
<th>Score of even numbered items</th>
<th>Difference of odd and even numbered items</th>
<th>$d^2$</th>
<th>$X_o + X_e$</th>
<th>$X_t$</th>
<th>$X_o - X_e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>48</td>
<td>69</td>
<td>-21</td>
<td>441</td>
<td>117</td>
<td>13689</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>102</td>
<td>94</td>
<td>8</td>
<td>64</td>
<td>192</td>
<td>38416</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>102</td>
<td>90</td>
<td>12</td>
<td>144</td>
<td>192</td>
<td>36864</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>77</td>
<td>89</td>
<td>-12</td>
<td>144</td>
<td>166</td>
<td>27556</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>98</td>
<td>71</td>
<td>27</td>
<td>729</td>
<td>169</td>
<td>28561</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>120</td>
<td>101</td>
<td>19</td>
<td>361</td>
<td>221</td>
<td>48841</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>92</td>
<td>79</td>
<td>13</td>
<td>169</td>
<td>171</td>
<td>29241</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>58</td>
<td>41</td>
<td>17</td>
<td>289</td>
<td>99</td>
<td>9601</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>63</td>
<td>48</td>
<td>-15</td>
<td>225</td>
<td>111</td>
<td>12321</td>
<td></td>
</tr>
</tbody>
</table>

contd.
Table 5.3 contd.

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>42</td>
<td>57</td>
<td>-15</td>
<td>225</td>
<td>99</td>
<td>9801</td>
</tr>
<tr>
<td>11</td>
<td>74</td>
<td>89</td>
<td>-15</td>
<td>225</td>
<td>163</td>
<td>28569</td>
</tr>
<tr>
<td>12</td>
<td>93</td>
<td>71</td>
<td>22</td>
<td>484</td>
<td>164</td>
<td>26896</td>
</tr>
<tr>
<td>13</td>
<td>56</td>
<td>38</td>
<td>18</td>
<td>324</td>
<td>94</td>
<td>8836</td>
</tr>
<tr>
<td>14</td>
<td>51</td>
<td>37</td>
<td>14</td>
<td>196</td>
<td>88</td>
<td>7744</td>
</tr>
<tr>
<td>15</td>
<td>58</td>
<td>71</td>
<td>-13</td>
<td>169</td>
<td>129</td>
<td>16641</td>
</tr>
<tr>
<td>16</td>
<td>107</td>
<td>0</td>
<td>0</td>
<td>214</td>
<td>45796</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>102</td>
<td>82</td>
<td>20</td>
<td>400</td>
<td>184</td>
<td>33856</td>
</tr>
<tr>
<td>18</td>
<td>27</td>
<td>29</td>
<td>-2</td>
<td>4</td>
<td>56</td>
<td>3136</td>
</tr>
<tr>
<td>19</td>
<td>54</td>
<td>71</td>
<td>-17</td>
<td>289</td>
<td>125</td>
<td>15625</td>
</tr>
<tr>
<td>20</td>
<td>54</td>
<td>39</td>
<td>15</td>
<td>225</td>
<td>93</td>
<td>8649</td>
</tr>
<tr>
<td>21</td>
<td>32</td>
<td>28</td>
<td>4</td>
<td>16</td>
<td>60</td>
<td>3600</td>
</tr>
<tr>
<td>22</td>
<td>66</td>
<td>79</td>
<td>-13</td>
<td>169</td>
<td>145</td>
<td>21025</td>
</tr>
<tr>
<td>23</td>
<td>103</td>
<td>89</td>
<td>14</td>
<td>196</td>
<td>192</td>
<td>36864</td>
</tr>
<tr>
<td>24</td>
<td>110</td>
<td>87</td>
<td>23</td>
<td>529</td>
<td>197</td>
<td>38809</td>
</tr>
<tr>
<td>25</td>
<td>52</td>
<td>44</td>
<td>8</td>
<td>64</td>
<td>96</td>
<td>9216</td>
</tr>
<tr>
<td>26</td>
<td>56</td>
<td>67</td>
<td>-11</td>
<td>121</td>
<td>123</td>
<td>15129</td>
</tr>
<tr>
<td>27</td>
<td>69</td>
<td>71</td>
<td>-2</td>
<td>4</td>
<td>140</td>
<td>19600</td>
</tr>
<tr>
<td>28</td>
<td>98</td>
<td>72</td>
<td>26</td>
<td>676</td>
<td>170</td>
<td>28900</td>
</tr>
<tr>
<td>29</td>
<td>56</td>
<td>43</td>
<td>13</td>
<td>169</td>
<td>99</td>
<td>9801</td>
</tr>
<tr>
<td>30</td>
<td>109</td>
<td>91</td>
<td>12</td>
<td>144</td>
<td>194</td>
<td>37636</td>
</tr>
<tr>
<td>31</td>
<td>96</td>
<td>74</td>
<td>22</td>
<td>484</td>
<td>170</td>
<td>28900</td>
</tr>
<tr>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>contd..</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>---</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>33</td>
<td>87</td>
<td>61</td>
<td>26</td>
<td>676</td>
<td>148</td>
<td>21904</td>
</tr>
<tr>
<td>34</td>
<td>63</td>
<td>78</td>
<td>-15</td>
<td>225</td>
<td>141</td>
<td>19881</td>
</tr>
<tr>
<td>35</td>
<td>96</td>
<td>83</td>
<td>13</td>
<td>169</td>
<td>179</td>
<td>32041</td>
</tr>
<tr>
<td>36</td>
<td>88</td>
<td>107</td>
<td>-19</td>
<td>361</td>
<td>195</td>
<td>38025</td>
</tr>
<tr>
<td>37</td>
<td>45</td>
<td>62</td>
<td>-17</td>
<td>289</td>
<td>107</td>
<td>11449</td>
</tr>
<tr>
<td>38</td>
<td>76</td>
<td>89</td>
<td>-13</td>
<td>169</td>
<td>165</td>
<td>27225</td>
</tr>
<tr>
<td>39</td>
<td>79</td>
<td>93</td>
<td>14</td>
<td>196</td>
<td>172</td>
<td>29584</td>
</tr>
<tr>
<td>40</td>
<td>85</td>
<td>61</td>
<td>24</td>
<td>576</td>
<td>146</td>
<td>21316</td>
</tr>
<tr>
<td>41</td>
<td>88</td>
<td>92</td>
<td>-4</td>
<td>16</td>
<td>160</td>
<td>32400</td>
</tr>
<tr>
<td>42</td>
<td>116</td>
<td>99</td>
<td>17</td>
<td>289</td>
<td>215</td>
<td>46225</td>
</tr>
<tr>
<td>43</td>
<td>112</td>
<td>97</td>
<td>15</td>
<td>225</td>
<td>209</td>
<td>43681</td>
</tr>
<tr>
<td>44</td>
<td>95</td>
<td>98</td>
<td>-3</td>
<td>9</td>
<td>193</td>
<td>37249</td>
</tr>
<tr>
<td>45</td>
<td>66</td>
<td>71</td>
<td>-5</td>
<td>25</td>
<td>137</td>
<td>18769</td>
</tr>
<tr>
<td>46</td>
<td>67</td>
<td>78</td>
<td>-11</td>
<td>121</td>
<td>145</td>
<td>21025</td>
</tr>
<tr>
<td>47</td>
<td>69</td>
<td>55</td>
<td>14</td>
<td>196</td>
<td>124</td>
<td>15376</td>
</tr>
<tr>
<td>48</td>
<td>41</td>
<td>32</td>
<td>9</td>
<td>81</td>
<td>73</td>
<td>5329</td>
</tr>
<tr>
<td>49</td>
<td>53</td>
<td>57</td>
<td>-4</td>
<td>16</td>
<td>110</td>
<td>12100</td>
</tr>
<tr>
<td>50</td>
<td>105</td>
<td>99</td>
<td>6</td>
<td>36</td>
<td>204</td>
<td>41616</td>
</tr>
<tr>
<td>51</td>
<td>115</td>
<td>97</td>
<td>18</td>
<td>324</td>
<td>212</td>
<td>44944</td>
</tr>
<tr>
<td>52</td>
<td>98</td>
<td>82</td>
<td>16</td>
<td>256</td>
<td>180</td>
<td>32400</td>
</tr>
<tr>
<td>53</td>
<td>19</td>
<td>28</td>
<td>-8</td>
<td>64</td>
<td>47</td>
<td>2209</td>
</tr>
<tr>
<td>54</td>
<td>95</td>
<td>108</td>
<td>-13</td>
<td>169</td>
<td>203</td>
<td>41209</td>
</tr>
<tr>
<td>55</td>
<td>111</td>
<td>127</td>
<td>-16</td>
<td>256</td>
<td>238</td>
<td>56634</td>
</tr>
</tbody>
</table>

contd.
<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>56</td>
<td>93</td>
<td>79</td>
<td>14</td>
<td>196</td>
<td>172</td>
<td>29584</td>
</tr>
<tr>
<td>57</td>
<td>12</td>
<td>23</td>
<td>-11</td>
<td>121</td>
<td>35</td>
<td>1225</td>
</tr>
<tr>
<td>58</td>
<td>97</td>
<td>88</td>
<td>9</td>
<td>81</td>
<td>185</td>
<td>34225</td>
</tr>
<tr>
<td>59</td>
<td>75</td>
<td>61</td>
<td>14</td>
<td>196</td>
<td>136</td>
<td>18496</td>
</tr>
<tr>
<td>60</td>
<td>87</td>
<td>62</td>
<td>25</td>
<td>625</td>
<td>149</td>
<td>22201</td>
</tr>
<tr>
<td>61</td>
<td>98</td>
<td>105</td>
<td>-7</td>
<td>49</td>
<td>203</td>
<td>41209</td>
</tr>
<tr>
<td>62</td>
<td>17</td>
<td>29</td>
<td>-12</td>
<td>144</td>
<td>46</td>
<td>2116</td>
</tr>
<tr>
<td>63</td>
<td>101</td>
<td>81</td>
<td>20</td>
<td>400</td>
<td>182</td>
<td>33124</td>
</tr>
<tr>
<td>64</td>
<td>104</td>
<td>91</td>
<td>13</td>
<td>169</td>
<td>195</td>
<td>38025</td>
</tr>
<tr>
<td>65</td>
<td>37</td>
<td>51</td>
<td>-14</td>
<td>196</td>
<td>88</td>
<td>7744</td>
</tr>
<tr>
<td>66</td>
<td>21</td>
<td>35</td>
<td>-14</td>
<td>196</td>
<td>56</td>
<td>3136</td>
</tr>
<tr>
<td>67</td>
<td>78</td>
<td>51</td>
<td>27</td>
<td>729</td>
<td>129</td>
<td>16641</td>
</tr>
<tr>
<td>68</td>
<td>99</td>
<td>81</td>
<td>18</td>
<td>324</td>
<td>180</td>
<td>32400</td>
</tr>
<tr>
<td>69</td>
<td>110</td>
<td>119</td>
<td>-9</td>
<td>81</td>
<td>229</td>
<td>52441</td>
</tr>
<tr>
<td>70</td>
<td>113</td>
<td>94</td>
<td>19</td>
<td>361</td>
<td>207</td>
<td>42849</td>
</tr>
<tr>
<td>71</td>
<td>27</td>
<td>55</td>
<td>-28</td>
<td>784</td>
<td>82</td>
<td>6724</td>
</tr>
<tr>
<td>72</td>
<td>59</td>
<td>72</td>
<td>-13</td>
<td>169</td>
<td>131</td>
<td>17161</td>
</tr>
<tr>
<td>73</td>
<td>96</td>
<td>75</td>
<td>21</td>
<td>441</td>
<td>171</td>
<td>29241</td>
</tr>
<tr>
<td>74</td>
<td>91</td>
<td>104</td>
<td>-13</td>
<td>169</td>
<td>195</td>
<td>38025</td>
</tr>
<tr>
<td>75</td>
<td>90</td>
<td>78</td>
<td>12</td>
<td>144</td>
<td>168</td>
<td>28224</td>
</tr>
<tr>
<td>76</td>
<td>41</td>
<td>49</td>
<td>-8</td>
<td>64</td>
<td>90</td>
<td>8100</td>
</tr>
<tr>
<td>77</td>
<td>78</td>
<td>59</td>
<td>19</td>
<td>361</td>
<td>137</td>
<td>18769</td>
</tr>
<tr>
<td>78</td>
<td>37</td>
<td>31</td>
<td>6</td>
<td>36</td>
<td>68</td>
<td>4624</td>
</tr>
</tbody>
</table>

Table 5.3 contd.
<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>79</td>
<td>56</td>
<td>71</td>
<td>-15</td>
<td>225</td>
<td>127</td>
<td>16129</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>107</td>
<td>118</td>
<td>-11</td>
<td>121</td>
<td>225</td>
<td>50625</td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>85</td>
<td>69</td>
<td>16</td>
<td>256</td>
<td>154</td>
<td>23716</td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>112</td>
<td>93</td>
<td>19</td>
<td>361</td>
<td>205</td>
<td>42025</td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>88</td>
<td>69</td>
<td>19</td>
<td>361</td>
<td>157</td>
<td>24649</td>
<td></td>
</tr>
<tr>
<td>84</td>
<td>108</td>
<td>115</td>
<td>-7</td>
<td>49</td>
<td>223</td>
<td>49729</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>111</td>
<td>93</td>
<td>18</td>
<td>324</td>
<td>204</td>
<td>41616</td>
<td></td>
</tr>
<tr>
<td>86</td>
<td>128</td>
<td>101</td>
<td>27</td>
<td>729</td>
<td>229</td>
<td>52441</td>
<td></td>
</tr>
<tr>
<td>87</td>
<td>87</td>
<td>61</td>
<td>26</td>
<td>676</td>
<td>148</td>
<td>21904</td>
<td></td>
</tr>
<tr>
<td>88</td>
<td>27</td>
<td>18</td>
<td>9</td>
<td>81</td>
<td>45</td>
<td>2025</td>
<td></td>
</tr>
<tr>
<td>89</td>
<td>68</td>
<td>75</td>
<td>-7</td>
<td>49</td>
<td>143</td>
<td>20449</td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>83</td>
<td>99</td>
<td>-16</td>
<td>256</td>
<td>182</td>
<td>33124</td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>103</td>
<td>113</td>
<td>-10</td>
<td>100</td>
<td>216</td>
<td>46656</td>
<td></td>
</tr>
<tr>
<td>92</td>
<td>97</td>
<td>73</td>
<td>24</td>
<td>576</td>
<td>170</td>
<td>28900</td>
<td></td>
</tr>
<tr>
<td>93</td>
<td>124</td>
<td>101</td>
<td>23</td>
<td>529</td>
<td>225</td>
<td>50625</td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>91</td>
<td>74</td>
<td>17</td>
<td>289</td>
<td>165</td>
<td>27225</td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>19</td>
<td>28</td>
<td>-9</td>
<td>81</td>
<td>47</td>
<td>2209</td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>29</td>
<td>37</td>
<td>-8</td>
<td>64</td>
<td>66</td>
<td>4356</td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>56</td>
<td>71</td>
<td>-15</td>
<td>225</td>
<td>127</td>
<td>16129</td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>74</td>
<td>63</td>
<td>11</td>
<td>121</td>
<td>137</td>
<td>18769</td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>98</td>
<td>103</td>
<td>-5</td>
<td>25</td>
<td>201</td>
<td>40401</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>37</td>
<td>49</td>
<td>-12</td>
<td>144</td>
<td>86</td>
<td>7396</td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>101</td>
<td>85</td>
<td>16</td>
<td>256</td>
<td>186</td>
<td>34596</td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>159</td>
<td>138</td>
<td>21</td>
<td>441</td>
<td>297</td>
<td>88209</td>
<td></td>
</tr>
</tbody>
</table>

contd.
<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>103</td>
<td>39</td>
<td>39</td>
<td>0</td>
<td>0</td>
<td>78</td>
<td>6084</td>
</tr>
<tr>
<td>104</td>
<td>97</td>
<td>85</td>
<td>12</td>
<td>144</td>
<td>192</td>
<td>33124</td>
</tr>
<tr>
<td>105</td>
<td>78</td>
<td>81</td>
<td>-3</td>
<td>9</td>
<td>159</td>
<td>25281</td>
</tr>
<tr>
<td>106</td>
<td>102</td>
<td>102</td>
<td>0</td>
<td>0</td>
<td>204</td>
<td>41616</td>
</tr>
<tr>
<td>107</td>
<td>106</td>
<td>99</td>
<td>7</td>
<td>49</td>
<td>205</td>
<td>42025</td>
</tr>
<tr>
<td>108</td>
<td>103</td>
<td>87</td>
<td>16</td>
<td>256</td>
<td>190</td>
<td>36100</td>
</tr>
<tr>
<td>109</td>
<td>97</td>
<td>81</td>
<td>16</td>
<td>256</td>
<td>178</td>
<td>31684</td>
</tr>
<tr>
<td>110</td>
<td>95</td>
<td>103</td>
<td>-8</td>
<td>64</td>
<td>196</td>
<td>39204</td>
</tr>
<tr>
<td>111</td>
<td>119</td>
<td>119</td>
<td>0</td>
<td>228</td>
<td>51984</td>
<td></td>
</tr>
<tr>
<td>112</td>
<td>71</td>
<td>83</td>
<td>-12</td>
<td>144</td>
<td>154</td>
<td>23716</td>
</tr>
<tr>
<td>113</td>
<td>28</td>
<td>28</td>
<td>0</td>
<td>56</td>
<td>3136</td>
<td></td>
</tr>
<tr>
<td>114</td>
<td>129</td>
<td>108</td>
<td>21</td>
<td>441</td>
<td>237</td>
<td>56169</td>
</tr>
<tr>
<td>115</td>
<td>106</td>
<td>85</td>
<td>21</td>
<td>441</td>
<td>191</td>
<td>36481</td>
</tr>
<tr>
<td>116</td>
<td>23</td>
<td>31</td>
<td>-8</td>
<td>64</td>
<td>54</td>
<td>2916</td>
</tr>
<tr>
<td>117</td>
<td>115</td>
<td>127</td>
<td>-12</td>
<td>144</td>
<td>242</td>
<td>58564</td>
</tr>
<tr>
<td>118</td>
<td>98</td>
<td>79</td>
<td>19</td>
<td>361</td>
<td>177</td>
<td>31329</td>
</tr>
<tr>
<td>119</td>
<td>95</td>
<td>77</td>
<td>18</td>
<td>324</td>
<td>172</td>
<td>29584</td>
</tr>
<tr>
<td>120</td>
<td>101</td>
<td>101</td>
<td>0</td>
<td>202</td>
<td>40804</td>
<td></td>
</tr>
<tr>
<td>121</td>
<td>56</td>
<td>71</td>
<td>-15</td>
<td>225</td>
<td>127</td>
<td>16129</td>
</tr>
<tr>
<td>122</td>
<td>59</td>
<td>38</td>
<td>21</td>
<td>441</td>
<td>97</td>
<td>9409</td>
</tr>
<tr>
<td>123</td>
<td>78</td>
<td>81</td>
<td>-3</td>
<td>9</td>
<td>159</td>
<td>25281</td>
</tr>
<tr>
<td>124</td>
<td>198</td>
<td>78</td>
<td>20</td>
<td>400</td>
<td>176</td>
<td>30976</td>
</tr>
<tr>
<td>125</td>
<td>99</td>
<td>81</td>
<td>18</td>
<td>324</td>
<td>180</td>
<td>32400</td>
</tr>
</tbody>
</table>

contd.
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>126</td>
<td>108</td>
<td>111</td>
<td>-3</td>
<td>9</td>
<td>219</td>
<td>47961</td>
<td></td>
</tr>
<tr>
<td>127</td>
<td>139</td>
<td>128</td>
<td>11</td>
<td>121</td>
<td>267</td>
<td>71289</td>
<td></td>
</tr>
<tr>
<td>128</td>
<td>86</td>
<td>71</td>
<td>15</td>
<td>225</td>
<td>157</td>
<td>24649</td>
<td></td>
</tr>
<tr>
<td>129</td>
<td>69</td>
<td>48</td>
<td>21</td>
<td>441</td>
<td>117</td>
<td>13689</td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>107</td>
<td>87</td>
<td>20</td>
<td>400</td>
<td>194</td>
<td>37636</td>
<td></td>
</tr>
<tr>
<td>131</td>
<td>102</td>
<td>105</td>
<td>-3</td>
<td>9</td>
<td>207</td>
<td>42849</td>
<td></td>
</tr>
<tr>
<td>132</td>
<td>98</td>
<td>82</td>
<td>16</td>
<td>256</td>
<td>180</td>
<td>32400</td>
<td></td>
</tr>
<tr>
<td>133</td>
<td>17</td>
<td>28</td>
<td>-11</td>
<td>121</td>
<td>45</td>
<td>2025</td>
<td></td>
</tr>
<tr>
<td>134</td>
<td>29</td>
<td>25</td>
<td>4</td>
<td>16</td>
<td>54</td>
<td>2916</td>
<td></td>
</tr>
<tr>
<td>135</td>
<td>96</td>
<td>79</td>
<td>17</td>
<td>289</td>
<td>175</td>
<td>30625</td>
<td></td>
</tr>
<tr>
<td>136</td>
<td>75</td>
<td>87</td>
<td>-12</td>
<td>144</td>
<td>162</td>
<td>26244</td>
<td></td>
</tr>
<tr>
<td>137</td>
<td>88</td>
<td>98</td>
<td>-10</td>
<td>100</td>
<td>186</td>
<td>34596</td>
<td></td>
</tr>
<tr>
<td>138</td>
<td>100</td>
<td>83</td>
<td>17</td>
<td>289</td>
<td>183</td>
<td>33489</td>
<td></td>
</tr>
<tr>
<td>139</td>
<td>71</td>
<td>61</td>
<td>10</td>
<td>100</td>
<td>132</td>
<td>17424</td>
<td></td>
</tr>
<tr>
<td>140</td>
<td>57</td>
<td>35</td>
<td>22</td>
<td>484</td>
<td>92</td>
<td>8464</td>
<td></td>
</tr>
<tr>
<td>141</td>
<td>78</td>
<td>84</td>
<td>-6</td>
<td>36</td>
<td>162</td>
<td>26244</td>
<td></td>
</tr>
<tr>
<td>142</td>
<td>96</td>
<td>83</td>
<td>13</td>
<td>169</td>
<td>179</td>
<td>32041</td>
<td></td>
</tr>
<tr>
<td>143</td>
<td>199</td>
<td>85</td>
<td>14</td>
<td>196</td>
<td>164</td>
<td>33856</td>
<td></td>
</tr>
<tr>
<td>144</td>
<td>108</td>
<td>87</td>
<td>21</td>
<td>441</td>
<td>195</td>
<td>38025</td>
<td></td>
</tr>
<tr>
<td>145</td>
<td>102</td>
<td>81</td>
<td>21</td>
<td>441</td>
<td>183</td>
<td>33489</td>
<td></td>
</tr>
<tr>
<td>146</td>
<td>69</td>
<td>69</td>
<td>0</td>
<td>0</td>
<td>138</td>
<td>19044</td>
<td></td>
</tr>
</tbody>
</table>
Table 5.3 contd.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>147</td>
<td>38</td>
<td>19</td>
<td>19</td>
<td>361</td>
<td>57</td>
<td>3249</td>
<td></td>
</tr>
<tr>
<td>148</td>
<td>105</td>
<td>93</td>
<td>12</td>
<td>144</td>
<td>198</td>
<td>39204</td>
<td></td>
</tr>
</tbody>
</table>

\[
\begin{array}{cccccc}
X_0 & X_e & d & d^2 & x_t & x_t^2 \\
11871 & 11107 & 764 & 33853 & 22978 & 3957120 \\
\end{array}
\]

Mean of old numbered items = 80.21
S.D. of old numbered items = 29.27
\( \sigma^2 \) of old numbered items = 856.73
Mean of even numbered items = 75.05
S.D. of even numbered items = 25.13
\( \sigma^2 \) of even numbered items = 631.52
\( \sigma_d \) = 14.95
\( \sigma^2_d \) = 223.50
Total Mean = 155.26
Total S.D. (\( \sigma_t \)) = 51.30
Total \( \sigma^2_t \) = 2631.69

The S.D. of the differences was found out and its value was 14.95. The S.D. of the whole test was 51.30. After finding out these figures the Eulon formula was applied to estimate the reliability.
The reliability is found to be .92 which is very high. This has led the investigator to say that the test is highly reliable and dependable.

5.2.5 Flanagan Formula

Flanagan's formula for establishing reliability is very close to Rulon's formula. In Flanagan's formula S.D. of two halves instead of difference of deviation of two sets are added. In other words, it estimates the error variance in a sense as the sum of the variance of the two halves. The formula is as follows:

\[ r_{tt} = 2 \left(1 - \frac{\sigma_1^2 + \sigma_2^2}{\sigma_t^2}\right) \]

Where \( \sigma_1^2 + \sigma_2^2 \) is used in the place of Rulon \( \sigma_d^2 \).

Where \( \sigma_1^2 \) = S.D. of scores of odd numbered items.

\( \sigma_2^2 \) = S.D. of scores of even numbered items.

\( \sigma_t^2 \) = S.D. of total score

---

The S.D. of two odds and even numbered item sets found out from the scattered diagram of split-half method in table 5.3. The S.Ds are 29.27 and 25.13 of odd and even numbered statements respectively. The S.D. for the total score is 51.30. These values were substituted in the formula as follows:

\[
rtt = 2 \left( 1 - \frac{\sigma^2 + \sigma^2}{\sigma^2_t} \right) = 2 \left( 1 - \frac{(29.27)^2 + (25.13)^2}{(51.30)^2} \right) = .87
\]

So the result is .87 which indicates that the test is highly reliable.

5.2.6 Method of Rational Equivalence

This method was developed by Kuder and Richardson. It is also known as K.R. Method. This method is useful for estimating internal consistency or homogeneity of the test. The Kuder-Richardson formula was developed because of the dissatisfaction with split-half method. A test can be split into two equal halves in great many ways and each split might yield some what different estimate of \( rtt \). The use of item statistics get away from such basis as may arise from arbitrary splitting into halves. Finally the most accurate
and practical formula was developed as follows:

\[ r_{tt} = \left( \frac{n}{n-1} \right) \left( \frac{\sigma^2}{\bar{t}^2} - pq \right) \]

Where

- \( n \) = Number of items in the test.
- \( p \) = Proportion of passing an item.
- \( q \) = \((1-p)\) Proportion of the group answering the item incorrectly.
- \( t \) = The S.D. of the test scores.
- \( r_{tt} \) = Reliability of the whole test.

This formula is called Kuder-Richardson's 20. Kuder-Richardson formula 20 is applicable only to test in which the items are scored by one point if answered correctly and nothing if not answered correctly. For the present CTA test, there are no answers of any items which could be treated in terms of right and wrong. Hence, the investigator decided to avoid this method.

5.2.7 The Analysis of Variance Approach

As so much of the statistical thinking is put in terms of variances, it is quite possible that the estimation of reliability can be made by a more convenient analysis of variance approach. Several investigators have proposed this

---

kind of approach. Among them are Jackson, Hoyt and Alexander. Like K-R approach, this one also starts with the item level.

Hoyt's basic formula of reliability is:

\[ r_{tt} = (1 - \frac{V_r}{V_e}) = \frac{V_e - V_r}{V_e} \]

Hoyt's regards the matrix of item scores as a two-way factorial design for analysis of variance without replications. According to Hoyt's formulation, variation in the performance of an individual from item to item is not considered to be error at all; rather it is a real (non-error) difference, an intra individual difference, and one which should be involved in the estimation of reliability. That is a total variation observed is conceived to be made up of three components: true-inter-individual differences, intra-individual differences (measured by item variance), and error inter-individual differences.

For finding out this, the data used in method of rational equivalence is used along with the data used in Rulon formula. The investigator has not used the method of rational equivalence because the items in the test are not indicating clearly yes-no category. Hence the method of analysis of variance is also not used by the investigator.

10. Ibid., p. 341.
5.3.0 Dependability of 'r'

In order to test the dependability of 'r' in terms of its SE, the researcher has to assume that the sampling distribution of 'r' is normal. The sample or population 'r' is placed at the centre of the distribution and SE of the 'r' as S.D. of the distribution. Since the probability is .01 of an 'r' exceeding $\pm 2.58$ there is only one chance in 100 that the error of $\pm SE \times 2.58$ or more is present in the 'r'.

If the obtained sample 'r' is in error, the researcher gets a serious error in estimating the population 'r'. If the sample distribution is relatively (smaller than 25) and 'r' is high, the sampling distribution of 'r' is skewed and the SE, obtained by using formula $(\frac{1-r^2}{N})$ is quite misleading. In such cases the skewness increases as 'r' increases. For the value close to $\pm .50$ and $N = 100$ or more the distribution of 'r' is supposed to be very close to the normal curve and formula $(\frac{1-r^2}{N})$ yields a useful estimate of significance.

Mathematically more defensible method of testing the significance of an 'r' especially when the co-efficient is very high or very low is to be converted in R.A. Fisher's function and find the SE of Z. The function has two advantages over 'r'. (1) Its sampling distribution is approximately normal. (2) Its SE depends only upon the size of the sample N, and is independent of the size of 'r'. The formula for $\sigma_2$ is as follows:
\[ \sigma_z = \frac{1}{\sqrt{N-3}} \]

(SE of Fisher's Z function)\(^{11}\)

From the Table C\(^{12}\) the investigator can read the 'z' corresponding to the obtained 'r' directly. The SE of Fisher's 'Z' by using formula \(\frac{1}{\sqrt{N-3}}\) is obtained. The confidence interval for the true 'Z' is \(Z \pm 2.58 \times \sigma_z\) for the .99 confidence interval range is obtained. The range of 'Z' again converted in 'r' range using the same table. This range of 'r' obtained by using Fisher's 'Z' technique is almost identical with range of 'r' obtained by \(SE_r\) if the sample distribution normal and adequately large. The range of \(SE_r\) and \(\sigma_z\) for .99 confidence interval level are given in table 5.4.

Table 5.4

<table>
<thead>
<tr>
<th>No.</th>
<th>Reliability Value</th>
<th>(SE_r)</th>
<th>Range of .99 confidence interval end</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Rest-Retest</td>
<td>.87</td>
<td>.02</td>
</tr>
<tr>
<td>2.</td>
<td>Split-half</td>
<td>.84</td>
<td>.024</td>
</tr>
<tr>
<td>3.</td>
<td>Rulon Formula</td>
<td>.92</td>
<td>.013</td>
</tr>
<tr>
<td>4.</td>
<td>Flanagan Formula</td>
<td>.87</td>
<td>.02</td>
</tr>
</tbody>
</table>

11. Ibid., p. 199.
12. Ibid., p. 448.
From the table 5.4 it is observed that $SE_r$ ranges obtained by using formula $\left( \frac{1-r^2}{N-3} \right)$ are almost identical. Hence it is concluded that the sample obtained for establishing reliability is quite normal and adequate. The reliability coefficients are quite high and dependable. Consequently the test is a quite reliable tool.

5.4.0 Comparison of the Reliability of the Present Test with Grome other Creative Thinking Ability Tests

After establishing the reliability of the test by various standard and popular methods, it is equally necessary to compare the present test with some other tests. The comparative study would make it clear whether the obtained reliability is good or poor. The table 5.5 shows the reliability coefficients of the present test and the reliability of other tests. The reliability coefficients of these scales were obtained from the research reports presented in a Survey of Research in Education Vol. I & II Edited by Buch, M.B.

<table>
<thead>
<tr>
<th>No.</th>
<th>Title of the test</th>
<th>Reliability co-efficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Passi Tests of Creativity</td>
<td>.92</td>
</tr>
<tr>
<td>2</td>
<td>Torrance's Test of Creative Thinking</td>
<td>.82</td>
</tr>
<tr>
<td>3</td>
<td>Baquer Mehdi Verbal Test of Creative Thinking</td>
<td>.78</td>
</tr>
<tr>
<td>4</td>
<td>Baquer Mehdi non-verbal Test of Creative Thinking</td>
<td>.81</td>
</tr>
<tr>
<td>5</td>
<td>Kaul Test of Creative Thinking</td>
<td>.88</td>
</tr>
</tbody>
</table>
The table 5.5 reveals that the reliability co-efficient of the present test is in no way less than that of the other tests. Nunnally Jr. Cum. G. has suggested that:

"In the construction of summative scales of attitude, usually 20 items selected in that way will have a reliability above .80."

This led the investigator to say that the reliability of the present test is quite high and highly satisfactory.

5.5.0 The Concept of Validity

A test is expected to prove its worth. If the test does not fulfil its worth, it is not worth anything. Ross and Stanley have rightly pointed out:

"Although high reliability is no guarantee that the test is good. Low reliability does not indicate that it is poor. Validity is always the first to be sought in a test, and granted that reliability is a valuable auxiliary."^{14}

---


The test for measuring mental abilities, aptitude, attitude etc., must justify their purpose. In the process of test construction and standardization, proof of justification of purpose is known as the scale validation. Consequently, validation of a test score is the most important and significant step in the process of standardization of any tests. Most of the users before selecting the test for use, look carefully into the values of validity. Therefore, the constructor of the test should make clear the concept of validity.

Validity is an important characteristic of a test. It depends upon the efficiency with which it measures what it attempts to measure. In other words, it is defined as the accuracy with which the test measures what it claims to measure. The term validity and purpose are very closely associated with each other. A test which fulfills the purpose for which it is designed is called a valid test. This means that a test of a creative thinking ability should measure creative thinking ability and nothing else. Therefore, in the course of a valid creative thinking ability test, the students who are creative should get more score than those who are less creative. Garrett has rightly put it as:

"The validity of the test or of any measuring instrument, depends upon the fidelity with which it measures what it purports to measure."  

Freeman has defined validity index as:

"An index of validity shows the degree to which a test measures what it purports to measure when compared with accepted criteria." 16

Freeman suggested that for validating a test, it must be compared with some accepted standards or other criteria which are regarded by experts as the evidence of the trait or ability to be measured by the test. Consequently the selection of validation criteria is of the prime importance in the process of a test validation.

5.6.0 The Methods Determining Validity

The test constructors often speak of operational validity and functional validity. Some constructors speak of empirical and rational validity. All these types of validities are discussed here. Procedures for determining validity of the tests are primarily concerned with the relationship between performance on the test and other independently observable facts about the behaviour characteristics under consideration. There are numerous techniques that are employed for investigating these relationships. Anastasi has presented them as follows:

"The APA Technical recommendations has classified these procedures under four categories, designated as content, predictive, concurrent and construct validity. Out of these four categories of validity the two, namely content and construct or concept validity are described under the headings of rational validity, by many authors. Similarly concurrent, predictive and congruent validity are described under the heading of empirical or statistical validity. In these methods the validity is estimated by means of statistical techniques."17

For the present test, the validity is estimated by measures of statistical techniques. The test constructor has established the following types of validity of the present test.

5.6.1 Construct or Concept Validity

In the case of determining the construct validity the first task is to define the measure. In the present test the creative thinking ability should be defined. Thorndike and Hegan have rightly explained what the phrase 'concept' or 'construct' really mean:

"Again we are thrown back on rational analysis, but this time we are trying to analyse concept and see what is

implied by it, rather than to make a catalogue of content."^{18}

For establishing the construct validity of the present test, the concept of creative thinking ability was translated into components. According to the term creativity it is analysed in the behaviour components. It is against this analysis the investigator should check his own test to see the construct validity. Thorndike and Hegan have rightly put it thus:

"The analysis and evaluation are now concerned not with content or subject matter acted upon but with the function of process that are applied to some content."^{19}

For establishing this type of validity, the investigator took the help of experts in the field of psychology, sociology and education, who know right type of evaluation process for analysing the statements of the test. After the careful analysis of the test statements describing the behaviour characteristics of a creative person, the basic characteristics of a creative person presented in the form of statements were approved by the experts. They are presented in table 5.6 along with their number in the scale and percentages.

19. Ibid., p. 113.
Table 5.6

**BASIC CHARACTERISTICS OF THE CREATIVE CHILD, STATEMENT'S NUMBER AND PERCENTAGES OF THE CHARACTERISTICS**

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Basic characteristics</th>
<th>Serial Number of the statements</th>
<th>Total</th>
<th>Percentages</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Shows capacity of discrimination</td>
<td>2,4</td>
<td>2</td>
<td>11%</td>
</tr>
<tr>
<td>2</td>
<td>Reports to logical and practicable imagination</td>
<td>1,3</td>
<td>2</td>
<td>11%</td>
</tr>
<tr>
<td>3</td>
<td>Knows proper comparison among things of familiarity</td>
<td>2,4</td>
<td>2</td>
<td>11%</td>
</tr>
<tr>
<td>4</td>
<td>Is capable of making proper judgements</td>
<td>3,5</td>
<td>2</td>
<td>11%</td>
</tr>
<tr>
<td>5</td>
<td>Sustains interest in maintaining relationship</td>
<td>2,4</td>
<td>2</td>
<td>11%</td>
</tr>
<tr>
<td>6</td>
<td>He has a more varied experiences of life situations</td>
<td>1,6</td>
<td>2</td>
<td>11%</td>
</tr>
<tr>
<td>7</td>
<td>Adopts proper reasoning in finding out merits and demerits of the things</td>
<td>5</td>
<td>1</td>
<td>5.5%</td>
</tr>
<tr>
<td>8</td>
<td>Is capable of finding out variegated uses of the things available</td>
<td>3</td>
<td>1</td>
<td>5.5%</td>
</tr>
<tr>
<td>9</td>
<td>Imagines and illustrates properly</td>
<td>4</td>
<td>1</td>
<td>5.5%</td>
</tr>
<tr>
<td>10</td>
<td>Does away with improper and unrelated things</td>
<td>2,3,4</td>
<td>3</td>
<td>17.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>18</td>
<td>100.0%</td>
</tr>
</tbody>
</table>
The statements are strictly in accordance with the behaviour characteristics. The test was presented to the judges for evaluating each statement keeping in view the behaviour characteristic. Thus the concept validity was established.

5.6.2 Congruent Validity

This type of validity is essentially estimated by the means of a statistical technique. For this the set of scores on the present test is correlated with the set of criteria of a similar measure. It means, it is correlated with some available well known powerful test of the similar nature. The correlation of the new test with the existing test would show to what extent the two scales measure the same behaviour characteristics. If the correlation coefficient is very high between these two sets, it is inferred that the new test is valid, since it measures what the criterion test is supposed to measure. The type of evidence just proposed is somewhat circular for this the condition is that the criterion test must be fully valid otherwise the correlation between the two tests would not be much dependable and this type of validity would be misleading. For finding out the congruent validity the investigator has used the Torrance Test of Creative Thinking Form B. The scores, obtained by 148 pupils on two different creativity tests were used as two sets of scores for finding out the correlation coefficient. The
correlation co-efficient was computed by the product moment method. The table 5.7 represents the distribution of the two sets of scores.

Table 5.7

THE SCATTER DIAGRAM OF SCORES OBTAINED BY THE STUDENTS ON C.T.A. TEST AND TORRANCE TEST OF CREATIVE THINKING

<table>
<thead>
<tr>
<th>Scores on Torrance Test</th>
<th>0-39</th>
<th>40-</th>
<th>80-</th>
<th>120-</th>
<th>160-</th>
<th>200-</th>
<th>240-</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>37</td>
<td>62</td>
<td>36</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>148</td>
<td></td>
</tr>
</tbody>
</table>

r = 0.82 ± 0.026

It is observed from the table 5.7 that the correlation co-efficient between the two sets of scores is 0.82 and SE_r is 0.026. This validity index is quite high and it provides an indication that the test is valid and dependable.

5.6.3 Concurrent Validity

The evidence of the validity is to be obtained from the relationship with other currently obtainable information
about an individual. Anastasi has defined the concurrent validity as follows:

"The relation between test scores and indices of criterion status obtained at approximately the same time is known as concurrent validity."\(^{20}\)

The test score in the investigation on hand means the score achieved in the C.T.A. Test by the respondents. The criterion status means the scores on the list of a hundred creative items rating scale. This rating scale is given in Appendix C. For establishing concurrent validity of the present test, technique of self-rating is used.

This tool is prepared by the investigator himself. The hundred items are so well-designed and thought out that each item has a specific value for measuring creativity, specifically prepared for the age-group 11 to 13. These items have been taken from different areas of creative activities of the said age-group. This test is purely a test based on children's own understanding and creative thinking. Items are verbal in their form and creative in their nature.

Its scoring is done by the total of the tick marks on the item-list. Item carries one mark each.

-------------------

This estimation about the creative thinking ability from the criterion stated above was compared with the C.T.A. scores. The obtained correlation is known as the co-efficient of concurrent validity.

The scores obtained by 148 pupils on C.T.A. test and self-done Activity Rating Scale were used as two sets of scores for finding out the correlation co-efficient. The correlation co-efficient was computed by the product moment method. The table 5.8 represents the distribution of the two sets of scores.

Table 5.8
THE SCATTER DIAGRAM OF SCORES OBTAINED BY THE STUDENTS ON C.T.A. TEST AND SELF DONE ACTIVITY RATING SCALE

<table>
<thead>
<tr>
<th>Self-Done Activity Rating Scale Scores</th>
<th>0- 1O- 2O- 3O- 4O- 5O- 6O- 7O- 8O- 9O- Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>36O-399</td>
<td>1 1</td>
</tr>
<tr>
<td>320-359</td>
<td>1 1</td>
</tr>
<tr>
<td>28O-319</td>
<td>1 1 2</td>
</tr>
<tr>
<td>240-279</td>
<td>1 1 1 3</td>
</tr>
<tr>
<td>200-239</td>
<td>1 4 3 3 3 1</td>
</tr>
<tr>
<td>160-199</td>
<td>3 3 6 3 1</td>
</tr>
<tr>
<td>120-159</td>
<td>5 9 5 3 1</td>
</tr>
<tr>
<td>80-119</td>
<td>17 11 8 1 1</td>
</tr>
<tr>
<td>40-79</td>
<td>2 27 16 1 2 1</td>
</tr>
<tr>
<td>0-39</td>
<td>1 2</td>
</tr>
<tr>
<td>Total</td>
<td>3 53 46 23 15 4 2 1 - 1</td>
</tr>
</tbody>
</table>

\[ r = 0.58 \pm 0.14 \]
It is observed from the table 5.8 that the correlation co-efficient between the two sets of scores is 0.58 and $SE_r$ is 0.14. This validity index is quite satisfactory and provides an indication that the test is valid and dependable.

5.7.0 Factorial Validity

It is a primary concern of the test constructor to find out the common functions of the measurement of the test. This entire test is called the common factor. For finding out this a certain type of statistical procedure is adopted which is called the factor analysis. Anastasi has explained it as follows:

"A major purpose of factor analysis is to simplify the description of behaviour by reducing the number of categories from an initial multiplicity of test variables to a few common factors or traits."\(^{21}\)

In order to isolate and study the factor composition of the test 210 answer-sheets were used.

5.7.1 Factorial Validity of the Test

For finding out the factorial validity of the test, the inter-correlation between six sub-tests were computed.
by product moment method. The correlation matrix of the above six tests is given below in table 5.9.

Table 5.9
CORRELATION MATRIX OF THE SIX TESTS

<table>
<thead>
<tr>
<th>Test</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(.28)</td>
<td>.04</td>
<td>.22</td>
<td>.08</td>
<td>.33</td>
<td>.28</td>
</tr>
<tr>
<td>2</td>
<td>.04</td>
<td>(.12)</td>
<td>.10</td>
<td>.06</td>
<td>.09</td>
<td>.12</td>
</tr>
<tr>
<td>3</td>
<td>.22</td>
<td>.10</td>
<td>(.52)</td>
<td>.52</td>
<td>.07</td>
<td>.07</td>
</tr>
<tr>
<td>4</td>
<td>.08</td>
<td>.06</td>
<td>.52</td>
<td>(.56)</td>
<td>.56</td>
<td>.07</td>
</tr>
<tr>
<td>5</td>
<td>.33</td>
<td>.09</td>
<td>.07</td>
<td>.56</td>
<td>(.64)</td>
<td>.64</td>
</tr>
<tr>
<td>6</td>
<td>.28</td>
<td>.12</td>
<td>.07</td>
<td>.07</td>
<td>.64</td>
<td>(.64)</td>
</tr>
</tbody>
</table>

5.7.2 Extraction of the Factors

The factor analysis of the present test was carried out by Thurstone's Centroid Method. The review of the above method is as follows:

This method has been developed on the basis of matrix algebra.

The term 'centroid' is borrowed from mechanics. It is a point in a mass where the centre of gravity is located. In factor analysis, the centroid of the end points of the test factors might be considered the locations of the centre of gravity of equal weight at the points. A
The purpose of factoring a correlation matrix is to account for the inter-correlations with fewer factors than there are tests. This factoring should be done so as to minimize the residuals after each factor has been determined. The main centroid axis is regarded as an approximation to the major principal axis of the factor configuration. This main centroid axis is so placed that it has zero projections on all the remaining co-ordinate axis. This leads to the rule that "the sum of the co-efficients in the correlation matrix is equal to the square of the sum of the first centroid loadings". The rule permits factoring through simple summational procedure after appropriate reflection. By 'Reflecting' it is meant that each of these factor retains its same length, but it extends in the opposite direction. The general policy is to reflect one test at a time and not the results, than to reflect a second one, and so on.

The extraction of each factor loadings reduces the residual in the correlation matrix. The factoring process is stopped when the standard deviation of the residuals is less than the standard error of a zero correlation.

With regards to sample size in Thurstone technique, Guilford advises to have minimum N of 200, when Pearson's
'r's are used. Hence factor loadings from sample here have been fairly consistent with loadings in the same factors and the tests.

Now for the present test data, the communalities of the test are not known. Therefore, they may be estimated from the data on hand. Fruchter has shown the simple method for this as follows:

"One simple method of estimating the communality of a test is to guess it to be equal to the highest correlation of that test with any other variable in the correlation table." 22

After deciding the communalities and entering them in row D the first factor was extracted as shown in table 5.10 and table 5.11 respectively.

Table 5.10

CENTROID FACTOR MATRIX SHOWING PROPORTION OF VARIANCE CONTRIBUTED BY THE CENTROID FACTORS AND COMMUNALITIES FOR THE C.T.A. TEST

<table>
<thead>
<tr>
<th>Test</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(.26)</td>
<td>.04</td>
<td>.22</td>
<td>.08</td>
<td>.33</td>
<td>.28</td>
<td>1.23</td>
</tr>
<tr>
<td>2</td>
<td>.04</td>
<td>(.12)</td>
<td>.10</td>
<td>.06</td>
<td>.09</td>
<td>.12</td>
<td>.53</td>
</tr>
<tr>
<td>3</td>
<td>.22</td>
<td>.10</td>
<td>(.52)</td>
<td>.52</td>
<td>.07</td>
<td>.07</td>
<td>1.50</td>
</tr>
<tr>
<td>4</td>
<td>.08</td>
<td>.06</td>
<td>.52</td>
<td>(.56)</td>
<td>.56</td>
<td>.07</td>
<td>1.85</td>
</tr>
<tr>
<td>5</td>
<td>.33</td>
<td>.09</td>
<td>.07</td>
<td>.56</td>
<td>(.64)</td>
<td>.64</td>
<td>2.33</td>
</tr>
<tr>
<td>6</td>
<td>.28</td>
<td>.12</td>
<td>.07</td>
<td>.07</td>
<td>.64</td>
<td>(.64)</td>
<td>1.82</td>
</tr>
<tr>
<td>E</td>
<td>1.23</td>
<td>.53</td>
<td>1.50</td>
<td>1.85</td>
<td>2.33</td>
<td>1.82</td>
<td>9.26</td>
</tr>
<tr>
<td>mE</td>
<td>.404</td>
<td>.174</td>
<td>.493</td>
<td>.608</td>
<td>.766</td>
<td>.598</td>
<td>3.043</td>
</tr>
</tbody>
</table>

$T = 9.26 \quad \sqrt{T} = 3.043$

$\frac{1}{\sqrt{T}} = .3286 = m \quad mT = 3.04$

The highest factor loading is .766 at the sub-test 5.
<table>
<thead>
<tr>
<th>$a_1$</th>
<th>Test 1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>.404</td>
<td>.117</td>
<td>-.030</td>
<td>.021</td>
<td>-.166</td>
<td>.021</td>
<td>.038</td>
</tr>
<tr>
<td>.174</td>
<td>-.030</td>
<td>.090</td>
<td>.014</td>
<td>-.046</td>
<td>-.043</td>
<td>.016</td>
</tr>
<tr>
<td>.493</td>
<td>.021</td>
<td>.014</td>
<td>.277</td>
<td>.220</td>
<td>-.308</td>
<td>-.225</td>
</tr>
<tr>
<td>.608</td>
<td>-.166</td>
<td>-.046</td>
<td>.220</td>
<td>.190</td>
<td>.094</td>
<td>-.294</td>
</tr>
<tr>
<td>.766</td>
<td>.021</td>
<td>-.043</td>
<td>-.308</td>
<td>.094</td>
<td>.053</td>
<td>.182</td>
</tr>
<tr>
<td>.598</td>
<td>.038</td>
<td>.016</td>
<td>-.225</td>
<td>-.294</td>
<td>.182</td>
<td>.282</td>
</tr>
<tr>
<td></td>
<td>.001</td>
<td>.001</td>
<td>.001</td>
<td>.002</td>
<td>.001</td>
<td>.001</td>
</tr>
<tr>
<td>E</td>
<td>.393</td>
<td>.239</td>
<td>1.065</td>
<td>1.019</td>
<td>1.701</td>
<td>1.057</td>
</tr>
<tr>
<td>mE</td>
<td>.186</td>
<td>.113</td>
<td>.505</td>
<td>.479</td>
<td>.333</td>
<td>.492</td>
</tr>
</tbody>
</table>

$T = 4.445 \quad \sqrt{T} = 2.108$

$\frac{1}{\sqrt{T}} = 0.4744 = m \quad mT = 2.11$

The highest factor loading is .505 of the sub-test 3.

After extracting every factor, criteria for sufficient factor were applied in order to stop extracting factors further. After extracting the second factor it was seen
by applying Humphrey's Rule whether the investigator should go further for extracting the factor. It was observed that the third factor would not be significant and, therefore, it was of no use to extract the third factor.

The Humphrey's Rules are reported hereafter.

5.7.3 Humphrey's Rule

This criterion takes into account the size of the sample and is dependent on loadings of only two variables rather than on entire matrix. The steps for the application of rule are as follows:

1. First find the product of the two highest loading in column of centroid matrix.

2. Find the standard error of a correlation coefficient of zero: for the type of correlation and size of sample being used. (e.g. \(1/\sqrt{N}\) for the Pearson Product moment of \(r\)).

3. If the product found in step 1 does not exceed twice the standard error found in step 2. The factor probably is not significant.
Table 5.12
SECOND RESIDUAL MATRIX WITH A COMPUTATION OF THE THIRD FACTOR LOADING OF THE C.T.A. TEST

<table>
<thead>
<tr>
<th>$a_2 = k_2$</th>
<th>Test 1</th>
<th>Test 2</th>
<th>Test 3</th>
<th>Test 4</th>
<th>Test 5</th>
<th>Test 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.186</td>
<td>0.082</td>
<td>-0.099</td>
<td>-0.073</td>
<td>0.077</td>
<td>-0.041</td>
<td>0.054</td>
</tr>
<tr>
<td>-0.113</td>
<td>0.009</td>
<td>0.077</td>
<td>-0.043</td>
<td>-0.008</td>
<td>0.005</td>
<td>-0.049</td>
</tr>
<tr>
<td>-0.505</td>
<td>0.073</td>
<td>-0.043</td>
<td>0.022</td>
<td>0.022</td>
<td>0.140</td>
<td>-0.023</td>
</tr>
<tr>
<td>-0.479</td>
<td>0.077</td>
<td>-0.008</td>
<td>-0.022</td>
<td>-0.039</td>
<td>-0.066</td>
<td>0.058</td>
</tr>
<tr>
<td>-0.333</td>
<td>-0.041</td>
<td>0.005</td>
<td>0.140</td>
<td>-0.066</td>
<td>-0.058</td>
<td>0.018</td>
</tr>
<tr>
<td>-0.492</td>
<td>-0.054</td>
<td>0.040</td>
<td>0.023</td>
<td>0.058</td>
<td>0.018</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td>0.002</td>
<td>0.001</td>
</tr>
</tbody>
</table>

\[ E = 0.336, \quad 1.82, \quad 0.323, \quad 0.270, \quad 0.328, \quad 0.233, \quad 1.672 \]

\[ mE = 0.260, \quad 0.141, \quad 0.250, \quad 0.209, \quad 0.254, \quad 0.160, \quad 1.294 \]

\[ T = 1.672, \quad \sqrt{T} = 1.293 \]

\[ \frac{1}{\sqrt{T}} = 0.7734 = m \quad mT = 1.29 \]

SE = 0.069, \quad 2(SE) = 0.138. Product of two highest factor loadings = 0.066

The product of the two highest factor loading of 0.069 does not exceed twice the value of SE of 0.064 (0.138). Consequently the third factor is not significant.
Now for this particular test with reference to table 5.12 it reads that the highest loading in column of the second factor are 0.260 (1) and 0.254 (5). The product of the two is 0.066 as the sample is of 210 the SE of the correlation co-efficient is \(1/\sqrt{210} = 0.069\). The value of the product of the two highest loadings 0.069 does not exceed twice the value of SE. Hence there is no need to go for the third factor, because it is not significant. Hence it could be said that there are only two significant factors which have proper loadings.

5.7.4 Interpretation of the Factors

In the previous paragraph it has been discussed that the third factor is not significant. It means that there are two common factors in all the six sub-tests. In order to find out how much percentages of total variance are taken away by these two factors the investigator calculated the contribution of the proportion of variance of these two factors and along with their communalities. The following table 5.13 shows the variance computation of these centroid factors.
Table 5.13

CENTROID FACTOR MATRIX SHOWING PROPORTIONS OF VARIANCE CONTRIBUTED BY THE CENTROID FACTORS AND COMMUNALITIES FOR THE C.T.A. TEST

<table>
<thead>
<tr>
<th>Test</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_1^2$</th>
<th>$a_2^2$</th>
<th>$h^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.404</td>
<td>.186</td>
<td>.163</td>
<td>.035</td>
<td>.198</td>
</tr>
<tr>
<td>2</td>
<td>.174</td>
<td>.113</td>
<td>.030</td>
<td>.013</td>
<td>.043</td>
</tr>
<tr>
<td>3</td>
<td>.493</td>
<td>.505</td>
<td>.243</td>
<td>.255</td>
<td>.498</td>
</tr>
<tr>
<td>4</td>
<td>.608</td>
<td>.479</td>
<td>.370</td>
<td>.229</td>
<td>.599</td>
</tr>
<tr>
<td>5</td>
<td>.766</td>
<td>.333</td>
<td>.587</td>
<td>.111</td>
<td>.698</td>
</tr>
<tr>
<td>6</td>
<td>.598</td>
<td>.492</td>
<td>.358</td>
<td>.242</td>
<td>.600</td>
</tr>
</tbody>
</table>

$\sum a_k^2 = 1.751$ $h^2 = 2.636$

Percentages

<table>
<thead>
<tr>
<th>$a_k^2$</th>
<th>$h^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>66.43</td>
<td>33.57</td>
</tr>
</tbody>
</table>

The distribution of the first factor is 66% and the distribution of the second factor is 34%.

It is observed from the table that the first factor has taken 64 per cent of the total variance and the second factor has taken 34 per cent of the total variance.

The first factor ($a_1$) has highest loading in sub-test no. 5 which measures the ability to see the deficiencies in the thing. It means it measures seeing defects
and seeing the needs, deficiencies, odds and the unusuals. Second factor \((a_2)\) has the highest loadings in the sub-test no. 3, which also measures the same type of ability of originality, novelty of description and novelty of imagination. The sub-test no. 1 measures the ability to forecast consequences. Sub-test no. 2 measures the ability to find similarities between the items. Sub-test no. 3 measures the ability to find out unusual uses of a given thing. Sub-test no. 4 measures ability to prepare lists of similar things. Sub-test no. 5 measures ability to see deficiencies in a given item and sub-test no. 6 measures ability of inquisitiveness. Similarly all the remaining tests measure some kind of creative thinking ability in one form or another. Therefore, the factor no. one \((a_1)\) and factor no. two \((a_2)\) run through all the sub-tests. This is a good indication of the test having a good factorial validity.

The present test through its common factors proves that it has a well established factorial validity of the entire test battery.