CHAPTER III

CONSTRUCTION AND STANDARDIZATION OF VOCATIONAL ANXIETY SCALE
Before making an attempt to prepare an instrument for measuring vocational anxiety, it becomes imperative to know its meaning. Again, we shall have to know the meaning and nature of general anxiety if we wish to understand the meaning of vocational anxiety.

More than four decades ago, Freud singled out anxiety as the crucial problem of emotional and behavior disorders. Further development of psycho-analysis has only substantiated his proposition, it is now recognised that anxiety is the fundamental phenomenon of neurosis or the dynamic centre of neurosis. Today, it is being felt everywhere that anxiety is much more prevalent than was suspected several decades ago. From the viewpoint of dynamic psychology, Symonds (1946) accurately notes that it would surprise most persons to realize how much of their behaviour is motivated by a desire to escape anxiety by either reducing it or disguising it in one way or another.

Whether we are concerned with normal or pathological behaviour, Freud (1920) was correct in saying that the solution to the riddle of anxiety must cast a flood of light upon our whole mental life.

A. MEANING OF ANXIETY

Giving an outline of the origin of the concept of anxiety, Wolpe (1972) writes, "Like its Latin original
anxietas, anxiety commonly connotes an experience of varying blends of uncertainty, agitation and dread. The Latin single usage included a suggestion of strangulation which is sometimes implied in the present-day connotation. The term was introduced in psychology when Freud (1894) described the anxiety neurosis as a syndrome distinct from neurasthenia. But its acceptance in the discipline did not become general until more than forty years later.

Mowrer (1936) views anxiety as a drive like other emotional tensions. The drive of an intense anxiety tends to dominate a person's behaviour. Freud (1936) conceived anxiety as an experiential phenomenon. It is a state that is perceived as basically unpleasant with physiological concomitants often present. Interpreting Freud's concept of anxiety, Hall (1956) states that Freud differentiated three types of anxiety—reality or objective anxiety, neurotic anxiety, and moral anxiety. These three types do not differ among themselves in any qualitative way. They all have the single quality of being unpleasant. They differ only in respect to their sources. In reality anxiety, the source of the danger lies in the external world. In neurotic anxiety, the threat resides in an instinctual object—choice of the Id. In moral anxiety, the source of the threat is the conscience of the superego system. An anxiety state may have more than one source. It can be a blend of neurotic and objective anxiety, or of moral and objective anxiety, or of neurotic and moral anxiety. It can also be a
blend of all three.

Interpreting anxiety biologically, Goldstein's central thesis is that 'anxiety is the subjective experience of the organism in a catastrophic condition.' It is agreed by students of anxiety—Freud, Goldstein, Horney, to mention only three—that anxiety is a diffuse apprehension, and that the central difference between fear and anxiety is that fear is a reaction to a specific danger while anxiety is unspecific, vague and objectless. Anxiety and fear in the classic psychoanalytic framework are distinguished in terms of dangers stemming from within and outside the individual, respectively. Special characteristics of anxiety are the feelings of uncertainty and helplessness in the face of the danger. The nature of anxiety can be understood when we ask what is threatened in the experience which produces anxiety. The threat is to something in the 'core' or essence of the personality. May (1950) characterizes anxiety as "the apprehension cued off by a threat to some value which the individual holds essential to his existence as a personality," and Sullivan (1953) refers to it as the state of tension arising from the experience of disapproval in inter-personal relations.

The threat may be to physical or psychological life (death or loss of freedom) or it may be to some other value which the individual identifies with his existence (patriotism, love of another person, success etc.). The occasion of anxiety will vary with different people because they hold different values of life. In anxiety, the threat is always to a value held.
by an individual as very essential to his existence and consequently to his security as a personality. An individual experiences various fears on the basis of a "security pattern" he has developed and in anxiety it is the 'security pattern' itself which is threatened.

B. VOCATIONAL ANXIETY

We come across terms like 'test anxiety,' 'school anxiety,' 'separation anxiety,' 'situational anxiety' etc., etc., in educational literature. This implies that many investigators have explored the possibility of establishing a phenomenon called 'situational' or 'specific' anxiety apart from general anxiety. It is generally taken to be a form of anxiety focussed upon a particular type of life stress.

We have stated above that anxiety may be caused by a threat to a value held by an individual as very essential to his existence and consequently to his security as a personality. There is every likelihood that the 'value' may be closely associated with the situation in which an individual is placed at a particular time. The cause of anxiety may be present within the individual or it may exist in the situation in which he happens to be placed at a particular time. If an individual possesses perpetual feelings of insecurity, inadequacy, helplessness etc., then, anxiety has become a part and parcel of his disposition, that is, he is temperamentally anxious person. On the other hand, if an individual finds himself insecure, disturbed, inadequate etc. in a particular situation, then he is experiencing 'situational' or 'specific' anxiety.
A large number of investigations have suggested that the more directly related the contents of items on an Anxiety scale is to the situations in which subjects are to perform, the more useful is the measure of anxiety in showing interactions between scores on the scale and differential motivating instructions.

Situational anxiety disturbs perceptual organisation, lowers concentration, attention, memory, reasoning, conceptual thinking, co-ordination and integration; and it creates stereotyped responses, disorganised activity, increased overt-activity, distractibility, irritability, heightened insecurity and responses which interfere with adequate and effective performance.

Like other parts of the World, we in India too, are passing through very difficult times. Urbanisation, mechanisation, changing values of life, population explosion, unemployment, poverty and rigid family relations are some of the factors which have made it difficult for an ordinary citizen to lead his life happily and peacefully.

The West perhaps never believed in the philosophy of seeking pleasure in renunciation but India which did for quite a long time, is rapidly escaping its hold. The values of life have changed, materialistic tendencies have occupied the Indian mind rather more firmly than it was expected and anxiety which was largely the monopoly of the West during the past three-four decades has ensnared the Indian population too.
Educational system in the country has not grown with time. Unplanned expansion of higher education, absence of adequate guidance services and lack of proper screening at lower stages have further aggravated the distressing situation. A sizable number of university graduates every year face a distressing and unpleasant problem of selecting and entering a career. Finding all other doors closed, most of them unwillingly decide to join a Teachers’ Training College. Some of them really intend to become teachers in their later lives but there are many others who simply wish to remain busy for one more year and thus plan to get time to find some more suitable job. After making up their mind to join a Teachers’ Training class, they again face a problem which is more serious and disturbing. How to get admission to a Teachers’ Training Course? It is not a simple affair as it seems to be. The intake capacity of Teachers’ Training Colleges is very much limited but there has been an unprecedented rush for admission to these colleges during the past few years. The applicants have to compete with numerous other applicants and only a lucky few are able to get a berth in these colleges. The selected students find themselves in a ‘New Land’ with which they were not familiar in their previous institutions. They are required to study for longer hours and are subjected to strict and rigid discipline of secondary schools. They have to study altogether new subjects which they had not studied previously at any stage. Programmes of student teaching and Discussion lessons are generally irritating for an average student teacher and they add to his
difficulties considerably. Besides, fear of future unemployment and resultant low status in family and society, keeps haunting their minds. These are some of the factors which make the student teachers anxious in the situation in which they have placed themselves willingly or unwillingly, happily or reluctantly.

Shaffer and Shoben (1956) have beautifully described the attitudes and strivings of a person experiencing anxiety in these words—"The drive of an intense anxiety tends to dominate a person's behaviour. To a person in severe conflict, the reduction of his anxiety is the most powerful urge to the exclusion of the relatively weaker motivations which lead to most of the positive satisfactions of life. As a result, people with strong, unresolved conflicts often develop mechanisms that have the sole utility of allaying their anxiety, even at the expense of sacrificing richer ultimate rewards."

Thus, a student teacher experiencing pangs of anxiety will direct his energies towards the reduction of his anxiety and will not be left with sufficient energy to enrich his educational and vocational life. Therefore, it is essential to detect the student teachers who experience 'situational anxiety' after joining a Teachers' Training Course, so that appropriate corrective services are arranged in time to set right their disturbed mental equilibrium. If they are allowed to go undetected, the natural outcome will be wastage and stagnation in Teacher Education and loss of valuable manpower to society.
In order to detect such students, we require an instrument to measure the specific kind of anxiety experienced by teacher trainees during the period of their training. We have labelled this kind of anxiety as vocational anxiety. Hence, vocational anxiety implies the anxiety experienced by teacher trainees with reference to their present educational status and future professional or employment status.

0. CONSTRUCTION OF ITEMS

Being a teacher in a Teachers' Training College, the investigator was in touch with student teachers. He decided to observe the behaviour of student teachers in class-rooms and elsewhere to detect the maladjusted students. He also made it known to all the students that any student having some adjustment problem was welcome to seek guidance from him. In this way, the investigator held detailed discussions with about fifteen such students who were either detected by the investigator or who had come forward voluntarily to seek guidance from him. These student teachers were encouraged to discuss their problems and difficulties frankly. These discussions equipped the investigator adequately for the big task of item construction. While constructing items, all the symptoms of vocational anxiety were kept in mind. Due help and guidance was taken from the MPI, MMPI, Taylor's Manifest Anxiety Scale, Sinha Anxiety Scale and Dutt Personality Inventory. The items taken from these instruments were redrafted and changed suitably to make them fit items for the measurement of vocational anxiety. Most of the items were constructed by the investigator himself.
The investigator came to the conclusion after studying many reports on scale construction and standardization that the length of the test does not matter as much as the judicious choice of items, item-analysis etc. Therefore, it was decided to have a scale of moderate size.

VA scale thus prepared, contained sixty items, each with three response choices—'Yes', '?', and 'No'. Some students of the B.Ed. class were asked to go through the scale and point out the language difficulties, if any. Most of them were corrected and others which could not be changed were noted for verbal explanation during administration. It was decided not to disclose the purpose of the scale to the students because it would have tempted them to give responses not characteristic of their behaviour. Therefore, it was christened as VA scale (Appendix I).

V. ITEM-ANALYSIS

The sixty-itemed scale was got printed and administered to about 175 students of B.Ed. (Bachelor of Education) class. Out of the total lot, 150 scales, complete in all respects, were selected for item-analysis. An equal number of scales of both the sexes were included in the 150 scales selected for item-analysis.

For item-analysis, various indices like phi coefficient, U-L 27% method, biserial r, point biserial r and tetrachoric r are generally employed. Adam (1960) has given his comments
upon the relative importance of these indices. He says,
"As the U-L 27% and phi (φ) have the advantage of computational ease, method and reliability which is quite comparable to the more painstaking methods, the U-L 27% method or phi should be acceptable for most situations." He further says,
"This latter conclusion is in essential agreement with the findings of Humphreys and Flanagan who have investigated the use of these methods of item-analysis from several other aspects."

Thus, Adam has strongly recommended U-L 27% method. Based upon this method, Sumner has recommended a very useful and satisfactory method known as the method of upper and lower thirds. In this case, the upper one-third of the lot forms the upper group and lower one-third of the lot forms the lower group. The mid one-third is discarded. The index of discrimination of an item is calculated by the formula given below:

\[ I.D. = \frac{U-L}{N/3} \]

Where

- I.D. = Index of discrimination of a particular item.
- U = The number of testees in the upper group showing the presence of the variable being tested.
- L = The number of testees in the lower group showing the presence of the variable being tested.
- N = Total number of testees.

This method is decidedly better than the method of U-L 27% because in the case of this method, the middle 33% of the cases are discarded, whereas in the method of U-L 27%, the middle
46% of the cases are not included in the final item-analysis. The I.D. computed by this formula is more reliable than the I.D. computed by using the method of U-L 27% because the size of upper and lower groups is comparatively large and thus, only very good items can hope to register an adequate index of discrimination. In the case of U-L 27% method, even some weak items may manage to get an adequate index of discrimination. Any item hoping to be included in a test must have the power to discriminate between high and low scorers.

This method also suffers from a serious limitation. The upper group in this method comprises only those testees who have admitted the presence of the variable which is sought to be tested. On the other hand, the lower group includes those testees who have denied the presence of the variable. In this way, the testees whose responses lie in between total acceptance and denial are not included in any of the two groups. Therefore, the Sumner's formula was further modified a little. The index of discrimination of each item was calculated as under:

\[ I.D. = \frac{U-L}{N/3} \]

Where

- \( I.D. \) = Index of discrimination of an item.
- \( U \) = Total scores of all the members of the upper group on the item under consideration.
- \( L \) = Total scores of all the members of the lower group on the item under consideration.
- \( N/3 \) = Maximum possible scores of 1/3 members of the whole group on the item.
In the case of the present VA scale, the maximum possible scores of 50 individuals on an item can be 100 because a maximum score of 2 has been assigned to each item. A score of two was assigned to a response indicating presence of vocational anxiety and a score of zero for a response indicating absence of vocational anxiety in an individual. It was also decided arbitrarily to allot a score of one to a response which was non-committal, that is, which was not indicative of either presence or absence of vocational anxiety in our subjects. The weights allotted to each response are given below:

(a) For items 8, 28 and 47

- 'Yes' carries a weight of 0
- '?' carries a weight of 1
- 'No' carries a weight of 2

(b) For all other items

- 'Yes' carries a weight of 2
- '?' carries a weight of 1
- 'No' carries a weight of 0

The 150 scales selected for item analysis were scored by the weights given to each response as shown above. All the scored scales were arranged in a descending order. The top scoring 50 scales comprised the upper group, whereas low-scoring 50 scales comprised the lower group. The middle 50 scales were eliminated. The index of discrimination of each item, as calculated by the modified Sumner's formula, is given in Table 1.
<table>
<thead>
<tr>
<th>Item No.</th>
<th>Total Scores U Group</th>
<th>Total Scores L group</th>
<th>Index of discrimination (I.D.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>79</td>
<td>54</td>
<td>.25</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>1</td>
<td>.24</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>3</td>
<td>.12</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>10</td>
<td>.40</td>
</tr>
<tr>
<td>5</td>
<td>95</td>
<td>83</td>
<td>.12</td>
</tr>
<tr>
<td>6</td>
<td>71</td>
<td>26</td>
<td>.45</td>
</tr>
<tr>
<td>7</td>
<td>69</td>
<td>15</td>
<td>.54</td>
</tr>
<tr>
<td>8</td>
<td>31</td>
<td>7</td>
<td>.24</td>
</tr>
<tr>
<td>9</td>
<td>36</td>
<td>14</td>
<td>.22</td>
</tr>
<tr>
<td>10</td>
<td>74</td>
<td>46</td>
<td>.28</td>
</tr>
<tr>
<td>11</td>
<td>76</td>
<td>46</td>
<td>.30</td>
</tr>
<tr>
<td>12</td>
<td>90</td>
<td>55</td>
<td>.35</td>
</tr>
<tr>
<td>13</td>
<td>51</td>
<td>13</td>
<td>.38</td>
</tr>
<tr>
<td>14</td>
<td>75</td>
<td>69</td>
<td>.06</td>
</tr>
<tr>
<td>15</td>
<td>58</td>
<td>38</td>
<td>.20</td>
</tr>
<tr>
<td>16</td>
<td>74</td>
<td>16</td>
<td>.58</td>
</tr>
<tr>
<td>17</td>
<td>85</td>
<td>57</td>
<td>.28</td>
</tr>
<tr>
<td>18</td>
<td>80</td>
<td>50</td>
<td>.30</td>
</tr>
<tr>
<td>19</td>
<td>29</td>
<td>4</td>
<td>.25</td>
</tr>
<tr>
<td>20</td>
<td>64</td>
<td>31</td>
<td>.33</td>
</tr>
<tr>
<td>21</td>
<td>32</td>
<td>7</td>
<td>.25</td>
</tr>
<tr>
<td>22</td>
<td>57</td>
<td>16</td>
<td>.41</td>
</tr>
<tr>
<td>23</td>
<td>17</td>
<td>0</td>
<td>.17</td>
</tr>
<tr>
<td>24</td>
<td>13</td>
<td>2</td>
<td>.11</td>
</tr>
<tr>
<td>25</td>
<td>15</td>
<td>7</td>
<td>.08</td>
</tr>
<tr>
<td>26</td>
<td>33</td>
<td>3</td>
<td>.30</td>
</tr>
<tr>
<td>27</td>
<td>14</td>
<td>0</td>
<td>.14</td>
</tr>
<tr>
<td>28</td>
<td>9</td>
<td>4</td>
<td>.05</td>
</tr>
<tr>
<td>29</td>
<td>51</td>
<td>16</td>
<td>.35</td>
</tr>
<tr>
<td>30</td>
<td>82</td>
<td>61</td>
<td>.21</td>
</tr>
<tr>
<td>Item No.</td>
<td>Total Scores U Groups</td>
<td>Total Scores L Group</td>
<td>Index of Discrimination (I.D.)</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------</td>
<td>---------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>31</td>
<td>49</td>
<td>25</td>
<td>.24</td>
</tr>
<tr>
<td>32</td>
<td>25</td>
<td>4</td>
<td>.021</td>
</tr>
<tr>
<td>33</td>
<td>74</td>
<td>46</td>
<td>.28</td>
</tr>
<tr>
<td>34</td>
<td>57</td>
<td>29</td>
<td>.28</td>
</tr>
<tr>
<td>35</td>
<td>40</td>
<td>12</td>
<td>.28</td>
</tr>
<tr>
<td>36</td>
<td>12</td>
<td>0</td>
<td>.12</td>
</tr>
<tr>
<td>37</td>
<td>8</td>
<td>0</td>
<td>.08</td>
</tr>
<tr>
<td>38</td>
<td>68</td>
<td>20</td>
<td>.48</td>
</tr>
<tr>
<td>39</td>
<td>61</td>
<td>22</td>
<td>.39</td>
</tr>
<tr>
<td>40</td>
<td>54</td>
<td>17</td>
<td>.37</td>
</tr>
<tr>
<td>41</td>
<td>12</td>
<td>4</td>
<td>.08</td>
</tr>
<tr>
<td>42</td>
<td>22</td>
<td>6</td>
<td>.16</td>
</tr>
<tr>
<td>43</td>
<td>42</td>
<td>4</td>
<td>.38</td>
</tr>
<tr>
<td>44</td>
<td>44</td>
<td>10</td>
<td>.34</td>
</tr>
<tr>
<td>45</td>
<td>71</td>
<td>25</td>
<td>.46</td>
</tr>
<tr>
<td>46</td>
<td>78</td>
<td>56</td>
<td>.22</td>
</tr>
<tr>
<td>47</td>
<td>17</td>
<td>2</td>
<td>.15</td>
</tr>
<tr>
<td>48</td>
<td>17</td>
<td>0</td>
<td>.17</td>
</tr>
<tr>
<td>49</td>
<td>14</td>
<td>1</td>
<td>.13</td>
</tr>
<tr>
<td>50</td>
<td>29</td>
<td>5</td>
<td>.24</td>
</tr>
<tr>
<td>51</td>
<td>25</td>
<td>0</td>
<td>.25</td>
</tr>
<tr>
<td>52</td>
<td>60</td>
<td>34</td>
<td>.26</td>
</tr>
<tr>
<td>53</td>
<td>12</td>
<td>4</td>
<td>.08</td>
</tr>
<tr>
<td>54</td>
<td>48</td>
<td>12</td>
<td>.36</td>
</tr>
<tr>
<td>55</td>
<td>45</td>
<td>14</td>
<td>.31</td>
</tr>
<tr>
<td>56</td>
<td>45</td>
<td>12</td>
<td>.33</td>
</tr>
<tr>
<td>57</td>
<td>57</td>
<td>24</td>
<td>.33</td>
</tr>
<tr>
<td>58</td>
<td>42</td>
<td>7</td>
<td>.35</td>
</tr>
<tr>
<td>59</td>
<td>44</td>
<td>15</td>
<td>.29</td>
</tr>
<tr>
<td>60</td>
<td>56</td>
<td>6</td>
<td>.50</td>
</tr>
</tbody>
</table>

**E. Selection of Items**

An item will have little discriminative value if both high and low-scorers give the same response to it. An item of this sort will not deserve inclusion in a standardized
psychological test. How much I.D. a good item must have? Most investigators generally fix .15 to .20 as the minimum desirable value of an I.D. Hence, in the present study also, it was decided arbitrarily that an I.D. with a value less than .20 does not significantly discriminate between the upper and lower groups.

The following items were not considered to be fit items for inclusion in the final test, because in the case of all these items, the value of I.D. was less than .20 — 3, 5, 14, 23, 24, 25, 27, 28, 36, 37, 41, 42, 47, 48, 49, 53.

As a result of rejection of the above mentioned 16 items, 44 items were left to be included in the final form of the scale.

No other method of item-analysis was used as only those items have been retained which have high discriminative value. Moreover, the method used for item-analysis is certainly a very robust method of item-analysis, and therefore, it was felt unnecessary to try some other method.

EFFECT OF SPEED AND CHANCE ON INDICES

Do speed and chance have any effect upon the indices calculated during the item-analysis? Since VA scale is administered under 'power conditions', the chance success is almost nil. There is no time limit for giving responses to the scale items. Therefore, the factor of speed does not affect our indices.

F. RELIABILITY OF VA SCALE

CONCEPT OF RELIABILITY:— Reliability has to do with accuracy and precision of a measurement procedure. Indices of reliability
give an indication of the extent to which a particular measurement is consistent and reproducible. Thus, a reliable test tests what it does test.

Thorndike and Hagen (1961) opine that there are two ways in which we can express the reliability or precision of a set of measurements. One indicates the amount of variation in a set of repeated measurements of a single specimen. We shall call this the "Standard error of measurement," since it is the Standard deviation of the errors of measurement. With psychological data, we can rarely actually make a whole set of measurements. Often we are fortunate if we can get two scores for each individual. But from such pairs of measurements, it is possible to get an estimate of the scattering of scores that would be found if we had made repeated measurements. Secondly, reliable measurement also implies that the individual stays in about the same place in his group. We can designate it as a reliability coefficient. The more nearly the measures are produced the second time, the higher the correlation and the more reliable the test.

To sum up, a measure is reliable to the extent that an individual remains nearly the same in repeated measurements as represented by a low standard error of measurement or by a high reliability coefficient.

METHODS OF ESTABLISHING RELIABILITY

There are four main methods of establishing reliability of a measurement procedure.

i) Test-Retest method.
ii) Parallel test method.

iii) Sub-divided test method.

iv) Reliability estimated from item-statistics.

Test-retest method means repeated administration of the same testing procedure and correlation of the resulting scores. Parallel test method includes administration of two equivalent forms of the same test and correlation of the resulting scores. Sub-divided test method means sub-division of a single test into two presumably equivalent groups of items, each scored separately, and the correlation of the resulting two scores. The last method of estimating reliability from item-statistics means analysis of the variance among individual items and determination of the error variance therefrom.

It is very difficult to develop a parallel form of the scale under study. It is comparatively very easy and desirable too in the case of intelligence tests and achievement tests. But in the case of personality inventories, construction of parallel forms is neither advisable nor practicable.

Test-retest method too could not be employed due to practical difficulties. Therefore, the investigator resorted to the remaining two methods which are equally useful in pointing towards the internal consistency of the test. The remaining methods are:

(a) SUB-DIVIDED TEST METHOD

This method is widely used to find out the internal consistency of a test. There are a number of ways to divide a test into two halves. A simple procedure, which is often relied upon to give equivalent halves, is to put alternate items into two
half-tests, that is, to put all the odd-numbered items in one half-test and all the even-numbered items in the other. This procedure is more sensible than dividing the test into first half and second half since items of similar form, content or discriminative power are likely to be grouped together in a test.

The two half tests will have a good probability of constituting equivalent tests. This procedure divides the test in two halves only for scoring and not for administration. This means that a single test is given at a single sitting and with a single time limit. However, two separate scores are derived—one by scoring the odd-numbered items and one by scoring the even-numbered items. The correlation between these two scores provides a measure of the accuracy with which the test is measuring the individual. But the correlation thus computed is between two half-length tests. This value is not directly applicable to full length test which is the actual instrument prepared for use. In general, the larger the sample of a person's behaviour we have, the more reliable the measure will be. Where the two halves of the test, which gave the scores actually correlated, are equivalent, we can get an unbiased estimate of total test reliability from the correlation between the half tests. The estimate is given by the Spearman-Brown Prophecy formula.

The main objection to this procedure is that a split-half reliability coefficient becomes meaningless when a test is highly speeded. The present test does not impose any time limit, therefore, the objection in this case is over-ruled.
b) RELIABILITY ESTIMATES FROM ITEM STATISTICS

According to Kuder and Richardson (1939), the method of rational equivalence represents an attempt to get an estimate of the reliability of a test, free from the objections raised against the methods outlined above. This method stresses the inter-correlations of the items with the test as a whole. Four formulas, popularly known as Kuder-Richardson formulas, are based on this method.

The Kuder-Richardson methods make the same assumptions as for the use of the Spearman-Brown formula. These assumptions call for items of equal, or nearly equal difficulty and inter-correlation.

PROCEDURE

The 'Y' form of VA scale (Appendix II) was administered to 325 student teachers studying in three different colleges of education in Panjab. Complete VA scales were selected for determining reliability of the scale. These 300 scales included 150 scales of male student teachers and the remaining 150 were those of female student teachers. All the scales were scored according to the weights allotted to each response, which are given below:

(a) For item No. 6

'Yes' carries a weight of 0

'? ' carries a weight of 1

'No ' carries a weight of 2
(b) For all other items

- Yes * Carries a weight of 2
- * ? * Carries a weight of 1
- * No * Carries a weight of 0

The weights assigned to each response make it evident that items have been keyed in the direction of the scale label, that is, higher scores reflect greater vocational anxiety.

Scores on odd and even items were totalled separately. A scattergram was plotted with total scores on odd items as Y-variable and total score on even items as X-variable. Pearson product-moment r was computed which came out to be .88. This value of correlation is between two half-lengths of the test. In order to find out the reliability coefficient of the full length test, the spearman-brown prophecy formula was applied. This formula is given below:

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

( Garret, 1969 a )

Where

- \( r_{tt} \) = Reliability coefficient of the whole test.
- \( r_{1/2} \) = Reliability coefficient of the half-test found experimentally.

When the reliability coefficient of the half-test is .88, the reliability coefficient of the whole test by the formula mentioned above is .94 (\( \frac{2 \times .88}{1 + .88} \)).

The reliability coefficient was also computed by the following K-R formula:

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

(Guilford, 1956 a)
Where
\( \bar{p} \) and \( \bar{q} \) = average proportions of passing and failing examinees for each item respectively.

The value of \( \bar{p} \) was obtained by dividing the mean of the total scores by \( n \). The average \( q \) is \( 1 - p \). The reliability coefficient was computed as under:

\[
rtt = \frac{44}{43} \left( \frac{193.21 - 44 \times 0.67 \times 0.33}{193.21} \right) = .97
\]

The significance of all these coefficients of reliability obtained as above for the VA scale was tested by Fisher's Z function. Table C from Garret was consulted for converting \( r \) into \( z \) and then \( z \) back into \( r \). Fiduciary limits of the obtained coefficients are given in Table 2.

**TABLE 2**

<table>
<thead>
<tr>
<th>Method used</th>
<th>( r )</th>
<th>Fiduciary Limits at .05 level</th>
<th>Fiduciary Limits at .01 level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odd-Even</td>
<td>.94</td>
<td>.925 to .950</td>
<td>.920 to .955</td>
</tr>
<tr>
<td>KR Estimates</td>
<td>.97</td>
<td>.960 to .975</td>
<td>.955 to .980</td>
</tr>
</tbody>
</table>

**INTERPRETATION**

The width of the fiduciary limits is so narrow that the trustworthiness of \( r \)'s is beyond doubt. Even if we accept the smaller coefficient of reliability, that is, .94 (Odd-Even), still there are 95 chances out of 100 that \( r \) of .94 may fluctuate from .925 to .950 and 99 chances out of 100 that it may fluctuate from .920 to .955. The obtained coefficient of
reliability is certainly very high. This fact assures us about
the trustworthiness of VA scale.

**Testing of Null Hypothesis**

The null hypothesis 'that population r's are zero' was tested by the formula:

\[ t = r \sqrt{\frac{N-2}{1-r^2}} \]  

( Guilford, 1956b)

The computed t values are put in Table 3.

<table>
<thead>
<tr>
<th>Method used</th>
<th>r</th>
<th>t</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odd-Sven ( N=300 )</td>
<td>.94</td>
<td>47.56</td>
<td>Significant well beyond .01 level</td>
</tr>
<tr>
<td>KR ( N=300)</td>
<td>.97</td>
<td>68.88</td>
<td>significant well beyond .01 level</td>
</tr>
</tbody>
</table>

Thus, the null hypothesis that the population r's are zero is rejected as the 't' value in both the cases is much higher than 2.592, the .01 level of significance with 298 degrees of freedom.

G. **Validity of Vocational Anxiety Scale**

Validity has to do with the question of what test scores measure and what they will predict. A test is valid if it measures what it claims to measure. In the case of vocational anxiety scale, we are concerned with the question — Does it really measure vocational anxiety? The scale will be valid if it measures vocational anxiety only and does not attempt to measure anything else.
TYPES OF EVIDENCE OF VALIDITY

A number of psychologists have given two main types of evidence bearing on the validity of a test—rational and empirical. In rational validity, we encounter a wide range of testing situations in which appraisal of the validity of a measurement procedure depends primarily upon the rational analysis and professional judgement. The analysis may be of the topics and areas included in the test—its content. For this type of analysis we shall speak of content validity. The rational analysis may be of the activities and processes that correspond to a particular concept (such as scientific method) and we may then speak of concept or construct validity.

The second main type of evidence of validity is empirical and statistical. This type of evidence comes from the relationship of the instrument that we are studying to some other measure or fact. This 'other measure' or fact may be very closely similar to our test or it may be quite different. It may be obtained at about the same time our test is given, or it may not be available for a long time in the future. Thus, there can be three further categories of the empirical and statistical evidence of validity.

(a) **CONGRUENT VALIDITY**

It is obtained by correlating a test with an existing similar measure of the same function.

(b) **CONCURRENT VALIDITY**

It is obtained by relating the test to some other measure obtained at the same time.
(c) **PREDICTIVE VALIDITY**

It is obtained by relating the test with some criterion of performance or success that becomes available in the future and is quite different from the test itself.

**PROS AND CONS OF EACH METHOD**

(a) **Congruent Validity:** The method is widely used to determine the validity of a measurement procedure. Its use is very common in validation studies, but in our case, it was not possible to make use of this method because no other scale or test of vocational anxiety was available. Very high coefficients of reliability of the VA scale indirectly speak of its congruent validity.

(b) **Concurrent Validity:** No 'some other measure,' other than tests, was available which could be of practical use in collecting data over a large sample of 300.

(c) **Predictive Validity:** This is, by far, the best procedure to establish validity, but it becomes very difficult to procure a suitable 'criterion measure'. The evidence of the effectiveness of our prediction is found in the coefficient of correlation between the test score and the later measure. Again no suitable 'later measure,' different from the present one could be developed. In the case of vocational anxiety, the later measure could only be a systematic record of an individual's behaviour symptomatic of anxiety. It was not possible for the investigator to adopt this very long procedure of establishing validity by following up 300 cases individually or collectively.
Thus, the investigator was left with only content validity and construct validity. Both of them were established rationally through critical examination and professional judgement. This procedure is quite in harmony with the procedure adopted by most of the investigators. Validity in the case of scales and inventories is generally established rationally. Logical relevance is not necessarily a second choice to empirical relevance in every instance. The assumption that a verbalized behaviour test will evoke relevant behaviour because the described situations are representative of typical real situations, is quite reasonable because the validity of personality inventories is often taken to be axiomatic if relevance has been established logically.

PROCEDURE

In practice, establishing the content and the construct validity of a test are often closely interwoven. Thus, the same steering committee that judges the importance of the content of different items may undertake to translate the underlying concepts (often relating to particular processes to be appraised) into manageable aspects of behaviour for testing. The two types of judgements may be made at the same time and by the same people. They represent, after all, two closely related aspects of the rational design or appraisal of a test.

Ten psychologists having good standing in the field of measurement, were approached to go through the VA scale. Three of them, have themselves done extensive research in the field of anxiety. They have also constructed and standardised anxiety
scales to measure anxiety in adults, adolescents and children. The investigator also met them personally and explained to them the meaning of VA and administration procedure of the scale. He sought from them answers to the following questions:-

What does VA scale measure ?
What is the nature of the psychological variable being measured ?
Have the items been written in accordance with instructions ?
Is administration procedure all right ?

All the judges were perfectly satisfied with the relevance of the scale. After getting green signal from them, the scale was got printed in its final form.

**ESTABLISHING VALIDITY INDIRECTLY FROM OTHER STATISTICS**

**Index of Reliability**— The degree to which a test measures what it measures, may be called the index of reliability or intrinsic validity. This definition can also be stated in terms of how well the obtained scores measure the test's true score component. The intrinsic validity is indicated by the square root of its reliability. The formula for finding index of reliability is given below:-

$$r_{t_\infty} = \sqrt{r_{tt}} \quad (\text{Garret, 1971 b})$$

Where

- $r_{t_\infty} =$ Index of reliability.
- $r_{tt} =$ Coefficient of reliability

Since it is directly and closely related to reliability, the same conditions that affect reliability will also affect intrinsic validity. Table 4 shows the indices of reliability for the VA scale which speak of its intrinsic validity.
TABLE 4
INDICES OF RELIABILITY FOR THE VA SCALE (Y-FORM)

<table>
<thead>
<tr>
<th>Method used</th>
<th>$r_{tt}$</th>
<th>$r_{te}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>KR</td>
<td>.97</td>
<td>.98</td>
</tr>
<tr>
<td>Odd-Even</td>
<td>.94</td>
<td>.97</td>
</tr>
</tbody>
</table>

For validity, the test has to be reliable first, only a highly reliable test can be a valid test. We cannot think of validity of an unreliable test.

Index of reliability is very useful in determining relation of validity to length of the test. Guilford (1954) remarks, "The nearer the index of reliability is to 1.00, the less effect will lengthening of any extent have on validity".

The values of indices of reliability in the present case are very near to 1.00, so, there is no need of lengthening the scale.

It has become evident that the present scale of vocational anxiety is a very reliable and valid instrument to measure vocational anxiety in teacher trainees. It will get refined and perfected in due course of time when it will be used in some later studies.
REFERENCES


Ibid 1971 b Page 349


