

The Effect of the Number of Response Categories of Likert Anxiety Trait Scale on its Psychometric Characteristics

Aiman Freihat *

aiman.freihat@bau.edu.jo

AL-Balqa Applied University (Jordan)

Receipt date: 10/2/2019; Acceptance date: 16/5/2019; Publishing Date: 30/6/2019

Résumé : Cette étude vise à identifier l'effet numérique des catégories de traits d'anxiété sur l'échelle numérique sur ses caractéristiques psychométriques. choisis au hasard et hiérarchiquement. Les résultats concernant les coefficients discriminés ont également mis en évidence l'existence de différences statistiquement significatives au niveau de la fonction ($\alpha = 0,05$) dans les coefficients discriminés entre les échelles triple et quatraine en faveur du quatraine, et entre les échelles triple et quinaire en faveur si le quinaire. La mise à l'échelle quinaire montre une grande fiabilité par rapport à la mise à l'échelle triple et à la mise à l'échelle quatraine.

Mots clés: Anxiété; Échelle de Likert; Caractéristiques psychométriques; Catégories d'échelle

Abstract: This study aims at identifying the number effect of response scaling categories of anxiety trait scale on its psychometric characteristics The study is applied on a sample of (352) male and female undergraduate students of Yarmouk University who are registered in the first semester 2017/2018 and selected randomly and hierarchically. Results concerning discriminated coefficients have also highlighted the existence of statistically significant differences at the function level ($\alpha = 0.05$) in discriminated coefficients between the triple scaling and the quatraine scaling in favor of the quatraine, and between the triple scaling and the quinary scaling in favor if the quinary. The quinary scaling shows high reliability in comparison with the triple scaling and the quatraine scaling.

Keywords: Anxiety; Likert Scaling; Psychometric Characteristics; Scaling Categories

*corresponding author

1-Introduction

Scientists, thinkers and researchers in psychology focus on finding ways or methods to help in the psychological phenomena in science or natural phenomena quantification. Psychological measurement faces difficulties in reaching the accuracy that has been reached by materialistic or natural measurement because the psychological measurement is indirect since it does not measure the quality or phenomenon directly but it measures the behavior that follows it; it is incomplete because it does not measure the whole quality or trait but a sample of it.

Accordingly, measuring the real score of the trait or psychological quality cannot be attained without errors since it is a must to find errors in experimented scores attained from measurement, errors that increase the real score or decrease it. These errors can be increased in size or decrease according to its source and accuracy, of which the most important is the scale itself, and its ability in measuring what it is designed for in addition to testing and the experimental conditions (Lord, 1960).

The errors of the psychological scale may be considered one of the most common errors affecting the measurement accuracy and its score. One of the main errors that cannot be controlled by the researcher since he, to an extent, can control the sample's errors by increasing its size and its representation of the society. The sample's errors can never exist if the trait is measured for all society. While, according to the errors of the experimental conditions associated to the measurement, those can be reduced by fixing or isolating them experimentally or statistically (Anderson & Brokowsk, 1978). By contrast, the errors of the measurement itself cannot be reduced no matter the accuracy of the procedures in the measurement formation since it is difficult to make the behavior sample that the scale tests representing accurately each behavior linked to the trait or the quality aimed to be measured by the scale. In addition to the fact that the psychological scale does not measure the trait directly but through what it signals or indicates of behaviors because the psychological traits are not described by its materialistic existence but by its virtual formation (Aiken, 1988).

So far, a lot of psychological studies of the psychological measurement scientists are directed towards finding new techniques and determining the measuring characteristics which reduce the scale's errors if they are available in a high degree, especially if they are found in paragraphs of scale since the scale as indicated by "Anastasi" is a neat and organized set of stimuli or paragraphs prepared in an objective and standardized manner to measure a sample of the behavior to be measured (Anastasi, 1988)

Studying the effect of response alternatives number on the anxiety scale is based on Likert's triple, quatrain, quinary scaling. In these types of scaling one is required to answer whether agreeing or disagreeing with differential degrees like highly agree, average or low. Likert's method has been widely used in a lot of psychological measurement studies, some of which aim at tracking the effect of scale categories number on the psychometric characteristics of the scale. The results of these studies vary since some of them have reported that the increase in the scale categories number generally increases the scale's reliability like Master's (1973), and Lissitz and Green (1975) and Cichetti, Showlter and Tryer (1985) study. On the other hand, Matell and Jacoby (1971) illustrate the independence of validity and reliability from the scaling categories number. Other studies have dealt with the average or neutral category effect on the scale characteristics such as Ray's study (Ray, 1980).

Because Likert's method depends on the degree of agreement, in addition to the fact that the agreement ranges from simple to absolute, the response alternatives scaling number varies, since some researchers depend on the triple scaling while some depend on the quinary scaling and few of them depend on sextuple scaling and rarely the scaling is nonan or more. The choice of response alternatives scaling number apparently depends on the researcher's

opinion and his own estimation. According to Likert's method in the light of the psychometric characteristics of each norm, there is no study on the response scales' items and numbers, which determines the preference of one scaling in the psychological measurements formation.

Thus, the underlying problem of the research is what the researcher faces when forming anxiety measurements in accordance to Likert's method by choosing or determining appropriate response alternatives scaling of the sample of individuals on whom the study is applied. Therefore, one of the primary methodology problems that face the researchers when forming psychological scales is due to the kind of response alternatives scaling of the item affects the judgment content of the respondent on the scale items content. Accordingly, responses of one item may vary by the difference of its response scaling norm.

Researchers used to, perhaps in personality measurement, the fact that response scaling of items is a verbal scaling in spite of the apparent differences concerning the number of these scales specially when the scale is formed by self-report which takes into consideration the reported sentences norm as one of the most common norms in forming its items that require the sentence to be with two, three, four, five or more response scale alternatives.

Anxiety is considered one of the common neurotic illnesses. It is the illness of the age inherent to man. Linguistically, anxiety means disorder and annoyance. We say: "He is extremely worried." That is, he is not stable on one state. Also we say: "Worry causes a person to be worried." That is, he is worried. (Academic lexicon, 1985). In spite of the fact that some researchers do not distinguish between anxiety and fear and see that they are synonymous, some of them find that it is necessary to distinguish between them (Abdel-Ghani, 1996).

Anxiety is considered to be in its ordinary and average degrees a natural response of man towards stimulants and certain situations. Moreover, it is considered one of the self-defense mechanisms and a way of staying alive used by man in facing threats and dangers which he is exposed to and it does not become dangerous if it stays within its normal range. It maybe a motivating and moving force and a source of activity for man which enables him to stay alarmed to dangers, to face the obstacles, and to confront all the challenges that arise along the prolonged struggle for survival. In this respect, anxiety leads to a positive role in man's life and pushes him to perform his activity (Michael, 2005).

Spielberger's theory of anxiety which makes use of the preceding theories and has two concepts of anxiety and they are: Anxiety state (A-State) and Anxiety trait (A-Trait). The A-State indicates anxiety as an urgent or temporary emotional state of man which differs in terms of intensity and fluctuates from one time to another (Spielberger, 1972).

1.1- The Problem of the Study and its Importance

Psychologists and researchers have tried to form psychological scales lacking errors as much as possible and they have looked for developing the methods of its formation and determining its characteristics in attempt to reach to some accurate estimations of the trait which the scale is formed for. That is because quantity helps in determining the form of behavior or the psychological functions and knowing the rules that control man's behavior. In addition to the fact that the measurement process in its core is noticed through information limited by the number of psychological phenomenon. Then, these numbers or the rules we get function to interpret man's behavior, control it, predict it and serve the individual and society (Rabe', 1994).

The phenomenon of quantitative process or the psychological traits that the psychological scale aims at calculating depends to a large extent on the way of judging its items. So, the individuals' scores on the scale are calculated through their responses to the

stimulants or its items which are normally determined by certain scale. Furthermore, the psychometric characteristics of the scale itself or its items, which accuracy indicates measuring what it is designed for, are mainly calculated from certain responses of the sample's individuals that are largely affected by the way of judging or answering its items (Fishben & Ajzen, 1973). There should be some psychometric characteristics available for the psychological scales items which lead to reducing as much as possible the errors of measurement results (Freeman, 1962).

A lot of studies have been carried out, which aimed at finding new formulas or paragraphs for the formulation of psychological scales have been conducted aiming to find forms of new items to structure psychological scales items. Some of them dealt with the form of paragraphs, and others dealt with the nature of the paragraphs or their standard contents, while some of them addressed the alternatives or their scales because all the psychometric properties of the scale or its vertebrae depend on the accuracy of the responses or the method of answering the paragraphs, which are usually affected by the type of staging and multiplexing (Johnson & Dixon, 1984)

At the same time, there are a number of studies concerned with the number of scaling points for response items, but these studies have never tackled the scaling of anxiety scale within the limited knowledge of the researcher. That is why, it is expected that the results of this study contribute to the suitable choosing of response alternatives scaling norms of which the triple, the quatrain, and the quinary of anxiety scale.

By going back to the problem of this study, it seeks to answer the following questions:

- Does the assessment of reliability scale sign differ by the response categories of the item?
- Does the assessment of validity scale sign differ by the response categories number of the item?
- Does the assessment of the discriminated sign value of the scale items differ by its response categories number?

1.2- Scope of Study:

The study is confined to Spielberger's anxiety scale as a study tool which is translated to Arabic by Dr. Amatanus Michael and it consists of (20) items of the trait.

It is confined to the students of humanity faculties and students of scientific faculties of the undergraduate in Yarmouk University of the academic year 2017/2018.

The categories number of Likert's scaling: is the number of response categories to the item of anxiety scale which include the following three norms:

1. The triple scaling norm (high, average, low)
2. The quatrain scaling norm (very high, high, low, very low)
3. The quinary scaling norm (very high, high, average, low, very low).

1.3-Theoretical Frame and Review of Literature:

1.3.1. Theoretical Frame Determining a theoretical frame is considered one of the basic requirements in conducting a research because it provides the researcher with a vision of the theoretical concepts that help him support his argument in all of the research procedures and results analysis, perhaps in forming scales since the researcher is required to determine the theoretical concepts and to attribute them to the procedures that these concepts impose. Moreover, in the light of the theoretical frame, the researcher achieves more of the procedures of the research and interprets its results (Gronlund, 1981).

The researcher deals with defining Spielberger's anxiety trait scale used in this study by clarifying its most important uses and the method of application and correction with brief

definition of psychological educational measurement and highlighting the psychoanalytical measurement theories.

The phenomenon of man fear and anxiety is one that coexists in his life. It is considered one of the important traits deeply rooted in his personality which expresses the reality of his existence. The 20th century has been called the age of anxiety in spite of the fact that worries about fear and anxiety trace back to ancient human history since anxiety has not been figured as a widely spread and recognized case of illness a bit before the beginning of the current century. Freud has been the first to note the important role of anxiety in the self-theory and in treating psychometric and psycho-physiological disorders.

Anxiety is the main phenomenon and central problem of the neurosis illness (Nervosas). According to Freud, anxiety is a thing felt as a state or certain sad emotional condition of man that encompasses experimental, psychological and behavioral aspects. During the past 50 years, clinical studies of man anxiety have been revealed increasingly and regularly in the psychoanalytic and psychoneurotic literature. Before 1950, studies of man anxiety have relatively been very few (Spielberger, 1983). Currently, the term anxiety is used to refer to two correlated and logically different terms. In practice, it is used to describe an emotional state or distress. Furthermore, anxiety has been used to describe a sort of static individual differences as one characteristic of the self (Spielberger, 1983).

1.3.1.2 Defining the Anxiety Trait Measurement: The concept of anxiety trait has been first dealt with by Catell (1996), and it has been developed by Spielberger (1966, 1972, 1976, and 1979). Generally, there are several states of the personality (Spielberger, 1972) such as, the emotional state that arise at certain moments of time, certain levels of intensity, and the anxiety states are the attributes of stress, anger, nervousness, sorrow, and activating or stimulating the automatic nerve system. In spite of the fact that the personality states are transitional, they happen when being motivated by suitable motives. These states may continue by the passage of time or the continuity of the motivation. Personality characteristics are described by a set of terms that Atkinson call motives and the Campbell (1963) indicate as behavioral attitudes acquired (Spielberger, 1983).

1.3.1.3 The Uses of the Anxiety Trait Scale: The anxiety trait scale has been used widely in evaluating clinical anxiety, operations, psychological illnesses and nerve patients. Generally, psychological patients and patients with depression have got high score on this scale. Moreover, the anxiety trait scale is used to classify high school students, university students, and military forces employees in identifying anxiety problems and evaluating immediate and long term outcomes of psychological and drug treatment. The anxiety trait scale has proved its convenience in experimental and clinical research in diagnosing people with high levels of nervousness (Spielberger, 1983).

Though the anxiety trait has been developed to be applied to the adults of high school universities' students; it has only proved its convenience regarding the early school students. Besides, the scale has also been adjusted and modified in more than 30 language for clinical research (Spielberger, 1983).

1.3.1.4 Anxiety Trait Scale Application: The anxiety trait scale has been designed to be self-applied and it can be given individually or collectively. There are no time limits for the scale. Generally, university students need approximately or about (6) minutes to fill in each scale and persons of less education or people with emotional disorders need (10) minutes to fill in each scale. It is required from those who set the test to build a relationship with the respondents before applying the anxiety trait scale (Spielberger, 1983).

When answering the anxiety trait scale, the respondents are given instructions to point out how their general feeling of their frequent anxiety feelings assessment on the points scale as follows:

Triple scaling: (1) low (2) average (3) high

Quatrain scaling: (1) very low (2) Low (3) high (4) very high

Quinary scaling: (1) very low (2) low (3) average (4) high (5) very high

The anxiety trait scale can be given with multiple choice answers which allow the auto-correction and the analysis like scattering papers and statistical packages to correct balanced responses (Spielberger, 1983).

1.3.1.5 The Anxiety Trait Auto-Correction: The scores value on the triple scale range from (20) score as a minimum limit to (60) score as a maximum limit, on the quatrain scale ranging from (20) score as a minimum limit to (80) score as a maximum limit, and on the quinary scale ranging from (20) scores as a minimum limit to (100) score as a maximum limit. There are also two kinds of expressions: The first refers to high score as high anxiety and the scales of this type are as follows: Triple Scaling (1,2,3), quatrain scaling (1, 2, 3, 4), and quinary scaling (1, 2, 3, 4, 5).

1.3.1.6 Concurrent Validity (The link between the scale and the personality scale):

The availability of links between the anxiety trait scale and other personality scales is a proof of concurrent validity. Generally, the existence of a lot of links with emotional and psychological disorder scales is expected. The high scores of university students' anxiety trait are simultaneous with bigger problems in every aspect of adjustment which have important meaning. These links signal that students exposed to anxiety develop problems in several aspects. Thus, the anxiety trait scale seems to be an active tool to diagnose students who need or look for an aid in student medical centers and health services.

1.3.2. Review of Related Literature: Al-Hazaimeh (1994) has conducted a study, titled "Factorial Construct of Likert's Scale by the Function Scaling Points Number", on a sample of (1360) male and female students selected successively and randomly from schools. He uses the attitude scale towards Math taken from the attitudes scale towards Math and its teaching methodologies formed and developed by each of Abu-Zeina and Al-Keilani (1978). Four forms have been used in accordance with Likert's scales used in the study and they are (2, 3, 4, 5).

1.4. The findings of the study: There is a positive relationship between the number of scale categories and coefficients number. The triple scaling is the nearest to form a scale whose factorial construct agrees with its logical theoretical construct.

There is no difference in values and homogeneity indices as an attribute of the difference in the number of scale categories. There is no relation between the number of scale categories and the percentage of the remaining items in the scale. Al-Akkam's (1995) study aims at dealing with the effect of scaling categories number of the attitude scale on the physics subject for the second year secondary scientific class.

The scale consists of (23) items scaled into triple, quatrain, quinary, hexa and sextuple, from "absolutely agree" to "completely disagree" and in equal number (230) student for each scale. The scale has been experimented on a sample of (1150) students in the region of Irbid distributed on (32) section in (18) schools (9 of them for male schools and other 9 are for females).

The analysis results show that there is no effect of the categories number of Likert's scale on the coefficients number. Besides, there are no differences in the reliability coefficients of different scales. Also, the results of variance analysis point at the existence of statistical significance differences owing to the differences in the scale categories number. In

attempt to test the homogeneity variance of different scaling in the scale, F-Test has been used. The results conclude to the fact that there is no difference in different variances owing to the difference in the scaling categories number of attitude scale towards physics.

Al-Ghamidi's (2003) study aims at knowing the range of impact of multiple alternatives response to anxiety scale, by the differences in academic stage, on validity coefficient and its reliability. The study has implemented anxiety test list scale modeled by Charles Splirger and others that describes the Saudi environment. Four norms have been formed for the scale differ in the response alternatives number. They are (two alternatives, three alternatives, four alternatives and five alternatives). In addition, the study has used the anxiety trait scale modeled by Charles Billberger and others that describe the Arab environment as external criterion of concurrent validity. The sample consisted of (451) student of general education in Jeddah city selected by multiple stage hierarchy. The study has concluded that the psychometric trait of the scale are affected by the multiple response difference since the values or reliability and validity coefficients increase by the increase of the response alternative number of the scale. Moreover, the change of academic stage affects on validity coefficient because the validity coefficient of the scale increases by the increase of mental or cognitive growth and knowledge of students. The reliability and validity coefficients values rise when the alternatives response number of the scale increases especially when the students get a high score of mental maturity and have a great motivation to participate.

Abu AlSal's (2008) study aims at comparing between the verbal scaling of self-report method scales and assessment of others method scales. The researcher has set a number of arbitrary zero-hypotheses to compare between the two scales of each method and between methods in each of the ordinal scale and verbal scale. He has chosen a sample of (420) student of the secondary stage in the city of Damascus. He has prepared an achievement motivation scale for the students of this stage consisting of (33) items in two methods: The self-assessment method and the assessment of others for each method. There are two scale answering norms of the items: The ordinal scale (1, 2, 3, 4) and the verbal scale (always, often, sometimes, rarely, never) and the sample individuals of student have to answer the self-rating method with the two norms for scale answering of the items. The parents of these students also answer the method of rating scale for others with two norms of answering its items. The researcher adapted the frequent design of measurements after dividing the sample into two groups and whether this sample is from the students or their parents. The first group has to answer the ordinal scale first, then the verbal scale. In the light of the research results, the researcher concludes that the ordinal scale is better than the verbal scale, perhaps in the light of the discriminated power of the items, their validity and reliability coefficients. Furthermore, the ordinal number, to a certain extent, is better for its method of others assessment than its method of self-report whereas there is no differences between them on the verbal scale.

In Matalle & Jacoby (1971) study titled by: Is There an Ideal Number for the Response Categories Number on Likert's scale?

The purpose of the study is to answer the following two questions: (1) Is there an ideal number for the scale categories number? (2) What is the range of the scale categories number impact on the scale's reliability and validity?

To answer these two questions, the Allport-Vernon-Lidzey values scale has been implemented. It contains 60 items with 18 different structure in the number of scales which range from 2 to 17 descriptive scale. The study concludes that validity and reliability are separate from the scale categories number.

Master (1973) has studied the relation between the scale points number and the reliability and discriminated items; this is by using a questionnaire of aspects towards

educational programs, and aspects towards educational traditions. The scales (6, 5, 4, 3, 2, 7) have been used and the sample study consists of (412) students in St. Petersburg University who were studying in the last stages. He has asked each one to answer on the two scales. After analyzing the student's answers, results have revealed that reliability and item discriminated increase basically by the increase in the scale points number from 4 to 6 to 7. Whereas Aikin (1983) has used a scale of teachers' assessment consisting of 10 items and 6 forms different in scales categories number (2, 3, 4, 5, 6, 7). While the original style of the scale is quinary scaled to show the effect of scale categories number on each of the reliability, the variance and the validity of the scale since the 6 forms have been applied on (624) student of each (104) student for each method.

After data analysis, the results have pointed out the increase of the item answers arithmetic means by the increase of the scale categories number. The relation between the answers variance and the scale categories number was curvilinear. Whereas the intra-coherence coefficient of total scores has changed regularly by the increase of scale categories.

Chan (1991) has also conducted a study about the rising effects of responses ordering on Likert's scale. He has used a tool consisting of 5 quinary scale items with two methods; the first starts from the positive end and ends in the negative end (absolutely agree, agree, neutral, disagree, completely disagree) and the second method is contrary to the first since it starts with the negative end and ends with the positive end. The tool of the study has been applied on a sample of (120) female and male students of high school in Taiwan and the results have proved that there is no effect for order difference (negative, positive) for scale points on the constructive coefficient.

Kim (1998) has also conducted a study in Japan aiming at limiting the perfect number of answer alternatives for the psychological scale items of the self-report method in the light of scale validity sign and its reliability. The study has used ordinal scale for four norms of answering; the first norm consists of 3 ordinal scales (3, 2, 1) and the second norm consists of 5 ordinal scales (5, 4, 3, 2, 1) whereas the third norms consists of 7 ordinal scales (7, 6, 5, 4, 3, 2, 1) and the fourth scale consists of 9 scale (9, 8, 7, 6, 5, 4, 3, 2, 1). The scale consists of (9) items and the study has used the random group design which contains groups, the total of which is (618) individuals. The study culminates to the fact that validity coefficient scale is better in the quinary scaling, whereas reliability is better in the quinary, sextuple, none, and it is lower in the triple scaling.

Preston and Colman (2000) has conducted a study aiming at comparing (10) norms of scale items answering alternatives prepared by the self-report method. These norms range from 2 alternatives to 11 alternatives with ordinal scores in the light of validity, reliability, and discriminated signs of scale items. The study has used frequent measurements design on a sample consisting of (149) male and female students of the university students in the United States of America. The study has culminated to the fact that the best norm among these norms is these norms of the quinary, hexa, and sextuple scale whether in the validity or reliability coefficient or the discriminated strength of the items.

1.5Comment on the Review of Related Literature: The study's results have varied concerning the scale categories number on the psychometric characteristics of the scale in accordance to Likert's scale. Of these studies are those agreeing with the increase of the scale categories number since this enables the increase of scale reliability and gives the respondent more options for response. Whereas, other studies show that the increase of scale categories number above a certain level decreases the reliability. On the other hand, other studies have shown that the scale's reliability and validity are separate from the scale categories number specially when the aim is measuring different levels of the trait. Besides, the fact that it is

distinguished by a bigger reliability coefficient than the scale of less categories. The increase of the scale categories number more than that may lead to a little reliability increase.

The majority of the preceding studies are based on three answering norms alternatives for the multiple choice and three alternatives in the first norm and four alternatives in the second norm and five alternatives in the third norm. Thus, the current research has used all of these three norms.

2. Methodology and Procedures:

2.1-Community of the Study: The community of study consists of Yarmouk University students according to the registrar and acceptance administration of the university for the bachelor degree of the first semester 2017/2018. The number of registered students according to the faculty and sex:

Table 1: Distribution of Individuals in the community of Study According to Two Study Variables (Faculty, Sex)			
Sex Faculty	Female	Male	Total
Arts	464	460	924
Educational Sciences	244	168	412
Economics& Administration	277	1047	1324
Law	59	150	209
Information Technology	138	290	428
Engineering	3	18	21
Total	1185	2133	3318

2.2The Sample of the Study: The sample of the study consists of (352) student randomly and hierarchically selected from each faculty in Yarmouk University. A number of academic sections have been chosen. Other sections have been chosen from the faculty of Arts' courses: English Communication Skills 1, Methodologies of Scientific Research for the Students of Arts, Islamic Fiqh Basics, and national Education and from the Educational Sciences Faculty: Rehabilitation of The special Needs, Administration of Children Nurseries, and from the Faculty of Information Technology: Computer Skills 1, Calculus, Introduction to Information Theory. From the Faculty of Law: Nationality Laws of the Foreigners, Methods of Legal Investigation. From the Faculty of Economics: Firms Accounting, Managerial Accounting, Accounting Theory, and The Management of Knowledge, Partial Economy, Data Structure Management, and Pricing Policies. From the Faculty of Engineering: Introduction to Computer and Telecommunication. The sample of the research is illustrated below:

Table 2: Distribution of Sample's Individuals According to Two Study Variables (Faculty, Sex)			
Faculty	Sex		Total
	Female	Male	
Arts	49	49	98
Educational Sciences	26	18	44
Economics	29	110	139
Law	6	16	22
Information Technology	15	31	46
Engineering	1	2	3
Total	126	226	352

2.3 Procedures of Forming Arab Image for Anxiety Scale:

1. Scale Validity: Arabic Translations of Jordanian environment have been detected by the scale after being examined by two teachers of English. Besides, the translation has been shown to two referees who are Arabic language PhD holders at Al- Balqa' Applied University/Ajloun University College.

2. Scale Reliability: Two courses in the first semester have been chosen and they are: Jordan Tourism Geography and Man & Environment in which students are from all specialties on whom the scale has been applied. The number of students in each course is illustrated in the following table:

Table 3: Distribution of the Explored Sample on Its Specialties for Validity Scael Purposes					
Faculty	Jordan Tourism Geography		Man & Environment		Total
	Female	Male	Female	Male	
Arts	7	8	27	8	27
Educational Sciences	4	1	9	2	9
Economics	6	20	53	20	53
Law	0	2	7	4	7
Information Technology	1	0	4	1	4
Engineering	1	0	1	0	1
Total	19	31	101	35	101

The scale has been applied repeatedly on them after (17) days for the purposes of scale reliability assessment (test-retest) then according to the reliability coefficient of Chronbach's Alpha method. the results are as follows:

Table 4: Explored Sample Analysis Results			
Scaling Categories	Cronbach's Alpha Reliability Coeffecient	(Test-Retest) Coefficient	Reliability
Triple Scaling	0.79	0.75	
Quatrain Scaling	0.82	0.78	
Quinary Scaling	0.844	0.81	

Chronbach's Alpha formula has provided results ranging between 79 to 84. And frequency reliability coefficients range between 75 to 81 and these results are very similar to those provided on Michael's scale.

3. Concurrent Validity

For the purpose of exploring the concurrent validity scal, neurosis scale has been applied to the research sample in addition to the study scale. Then, the correlation coefficient has been calculated between the research sample individual responses on both scales. Here are the results as illustrated in the following table:

Table 5: Concurrent Validity Analysis Results Between The Two Scales	
Categories Scaling	Correlation Coefficient with Neurosis Scale
Triple Scaling	0.65
Quatrain Scaling	0.75
Quinary Scaling	0.82

By reading the above table, it is observed that the neurosis scale provides a high correlation with the anxiety trait scale which in turn hints at the strong relation between anxiety and neurotic trait and this supports the scale validity.

3- The Study Tools:

1. Spielberger's Anxiety Scale

T. Anxiety Scale has been implemented as anxiety trait scale which consists of (20) decisive phrases aiming at assessing the feelings of the person being tested in general. The modified image which appeared in (1938) also has been used showing several modified images in a number of Arab countries. The work on Spielberger's scale has started in 1964.

The researcher has based his study on the modified image of the anxiety scale more than the original image in 1983 which used to be in several studies in Arab countries (Michael, 2006) since three images have been formed:

1. Likert Triple Scaling with the score (high, average, low).
2. Likert Quatrain Scaling with the score (very high, high, low, very low).
3. Likert Quinary Scaling with the score (very high, high, average, low, very low).

3.1 Statistics Treatment: After gathering data, the statistics treatment has completed by using the statistics analysis program SPSS since the required statistics have been calculated like Cronbach's Alpha coefficient, correlation coefficient of item-correction, finding the mean single analysis of variance for the discriminated values of items ANOVA, and statistical M, Z, and U and have been calculated for correlated samples.

3.2 The findings related to the first question: Does the reliability scale sign assessment differ depending on the response categories number for each item?

In answering this question, the reliability coefficient has been calculated by using Croonbach's Alpha-formula for each form of the three scales. Illustrated on table 6:

Table 6: Value of Reliability Coefficient Using Chronbach Formula For Each Form of The Scale	
Scaling Categories	Cronbach's Alpha Reliability Coefficient
Triple Scaling	0.7143
Quatrain Scaling	0.8246
Quinary Scaling	0.8506

In looking to table (6), it is noticed that results show that the value of the reliability coefficient for the quinary scale (.8506) is higher than the value of reliability coefficient for the triple scale (.7143). These are high and acceptable values expressing the quality of reliability of the anxiety trait scale for the study sample individuals in comparison to Michael's scale. For revealing the difference in the coefficient values with statistical significances among the three forms of the scale, then the value M has been found, rule (Hasktain & Whalen, 1976) in the following formula:

$$M = \frac{J - 1}{18J} \left[\sum_{k=1}^3 B_k - \frac{\left[\sum_{k=1}^3 B_k (1 - \alpha_k)^{-1/3} \right]^2}{\sum_{k=1}^3 B_k (1 - \alpha_k)^{-2/3}} \right]$$

The value of B is calculated:

$$B_k = \frac{(9n_k - 1)^2}{(n_k - 1)}$$

* **Symbol K refers to: form number**

NK: The number of sample individuals in the form.

AK: Chronbach-Alpha Reliability Coefficient.

The value M has reached (35.94446) and it has a statistical significance at the significance level ($\alpha = 0.05$). This means that differences of statistical significance exist between reliability coefficients values when compared at the critical level K2 (5.99) with two free degrees and for detecting the differences between any couple of coefficient. They have been tested by transferring them to Z-Fisher value. Then, using statistical Z to reveal the significance of differences between every two reliability coefficients by using the following formula:

$$Z = \frac{z_1 - z_2}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

* **Symbols Z1, Z2 refer to the value of Z-Fisher opposite to each of n1, n2 the number of individuals of each sample.**

The results of the comparison are illustrated on table 7:

Table 7: Significance of Differences Between The Reliability Coefficients , Chronbach's Alpha and Statistical Value Z of these differences (N= 352)	
Scaling Categories	Z value
Triple with Quatrain	3,567 *
Triple with Quinary	4,874 *
Quatrain with Quinary	1,308 -

* **Function at the significance level ($\alpha = 0.05$)**

In studying table 7, the results have shown by comparing the calculated values with Z critical value $z=1.96$ at the significance level we find z more than 1.96. Consequently, the sign between the triple scale and quatrain scale is in favor of the quatrain; the sign between the triple and quinary is in favor of the quinary, and another sign between the quatrain scale and the quinary.

3.3 Results Related to the Second Question: Does the assessment of concurrent validity sign of the scale differ by the number of responses for the item?

In answering this question, the correlation coefficient has been calculated as a sign for concurrent validity of the anxiety trait scale which represents a tool for this study. By the calculation of concurrent correlation coefficient of the explored sample individuals among the response of sample individuals on the anxiety trait scale and neurosis scale.

Table 8: The Value of Concurrent Validity Coefficient for the Three Forms of the Study Tool and the Statistical Value (V)

Scaling Categories	Correlation Coefficient between Neurosis Scale and The Study Tool	Z-Fisher Value
Triple Scaling	0.65	0.0775
Quatrain Scaling	0.75	0.0973
Quinary Scaling	0.82	1.157
V value	*7.153	

*** Function at the significance Level $\alpha=0.05$**

since:

$$I = \sum_{j=1}^3 \left(n_j - 3 \right) \left(z_j - u \right)^2$$

u value is calculated as follows:

$$u = \frac{\sum_{j=1}^3 \left(n_j - 3 \right) z_j}{\sum_{j=1}^3 \left(n_j - 3 \right)}$$

n_j: number of sample individuals in form j.

z_j: The Fisher mark opposite to correlation coefficient j

j: form number

As illustrated on table 8, the results show that there are differences of statistical significance at the significance level $\alpha=0.05$ between the concurrent validity coefficients of the anxiety trait scale and neurosis scale, since the value of v is calculated (7.153) and it is a statistical function at the significance level $\alpha=0.05$. It is considered a function when compared in K2 critical value (5.99) with free degree Z1 and for detecting the differences, they have been tested by the Z-Fisher test to reveal the differences between the concurrent validity coefficients as shown on table 9:

Table 9: Significance of differences Between the Concurrent Validity Coefficients and Statistical Value Z of these Differences

Scaling Categories	Z Value
Triple with Quatrain	1.386
Triple with Quinary	-*2.671
Quatrain with Quinary	1.287

*** Function at the Significance level $\alpha=0.05$**

Table 9 indicates that there are differences of statistical function between the concurrent validity coefficient vales at the significance level ($\alpha=0.05$) between the triple scale and the quinary scale in favor of the quinary owing to the difference in scale categories; whereas, no differences of statistical significance have appeared between the concurrent validity coefficients values owing to the difference in scale categories between the triple and the quatrain scale, and between the quatrain and the quinary scale.

3.4 Results related to Question three: Does the discriminated sign assessment of the scale items differ by the categorie response number?

In answering this question, the discriminated coefficients of the items in each of the three styles have been calculated by finding the corrected correlated coefficient item with the

scale and has its three styles as a sign of the scale's validity. Moreover, the means and standard deviation of the coefficients have been calculated. The values of these coefficients for the triple scale have reached between (.1 and .5) with a mean value (.28) and standard deviation (.97) and for the quatrain scale between (.3 and .6) with a mean value (.45) and standard deviation (.94). These are accepted values on the scale and accredited as a sign of the constructive validity signs of anxiety trait scale in accordance to Michael's scale results.

Table 10: Corrected Correlation Coefficient Values of the Item For Each Style of the Scale

Triple Scaling	Corrected Correlation Coefficient	Quatrain Scaling	Corrected Correlation Coefficient	Quinary Scaling	Corrected Correlation Coefficient
A1	0.1985	B1	0.2984	C1	0.3153
A2	0.1497	B2	0.2442	C2	0.3285
A3	0.2718	B3	0.4929	C3	0.5048
A4	0.2072	B4	0.2673	C4	0.3482
A5	0.3691	B5	0.4354	C5	0.4949
A6	0.4323	B6	0.5512	C6	0.5835
A7	0.1767	B7	0.3410	C7	0.3847
A8	0.2129	B8	0.4690	C8	0.5306
A9	0.3322	B9	0.4867	C9	0.5491
A10	0.4571	B10	0.4950	C10	0.5521
A11	0.0631	B11	0.4808	C11	0.5374
A12	0.3578	B12	0.3348	C12	0.3953
A13	0.3513	B13	0.5004	C13	0.5202
A14	0.3287	B14	0.5074	C14	0.3368
A15	0.3264	B15	0.3338	C15	0.4354
A16	0.3910	B16	0.2941	C16	.3692
A17	0.3649	B17	0.4141	C17	0.4928
A18	0.2206	B18	0.4514	C18	0.5168
A19	0.2141	B19	0.2869	C19	0.3684
A20	0.2995	B20	0.3620	C20	0.4177
Mean	0.2821		0.4042		0.4542
Standard Deviation	0.09964		0.09509		0.9380

As illustrated on table 10, the results have shown that correlation coefficient item score of the quinary scale is higher than that of the quatrain scale. Moreover, the correlation coefficient item score of the quatrain scale is higher than that of the triple scale with big differences. The mean value of the triple scale form has reached (.2128) and a standard deviation (.64099) and the mean value of the quatrain scale form (.4042) and standard deviation (.09509) and the mean value of the quinary scale form (.4245) and the standard deviation (.09380). It is clear from the table itself that the quinary scale is the highest in the arithmetic mean of discriminated coefficient, next the quatrain scale and then the triple scale. In order to know the difference in means of discriminated coefficients among the three styles, a variance analysis of the frequent scales has been used to reveal the differences in the means of discriminated coefficients among the three styles and the results are illustrated in table 11.

Table 11: Results of Variance Analysis of The Frequent Measurements of The Study Sample Individuals On The Three Styles of The Scale

Variance Source	Degrees of Freedom	Total of Squares	Means of Squares	F-Value	Statistical Significance
Between Individuals	19	0.3859	0.203		
Between Measurements	2	0.3136	0.1568	*42.0296	0.000
Error	38	0.1417	0.0037		
Total	59	0.8412	0.0143		

* **Function at the significance level $\alpha=0.05$**

The above table shows the existence of an effect for the scale categories number difference on the mean of discriminated coefficients in the three styles. Since the value f indicates the existence of differences of statistical significance at the significance level ($\alpha=0.05$) for ($\alpha=0.05$, $F=42.0296$). Consequently, there is an effect for the difference in the number of alternatives on the mean of the discriminated coefficients in the three styles. That's why an analysis has been done to clarify the differences in table 12:

Table 12: Results of Discriminated Analysis Values of Frequent Measurements For The Sample Study Individuals to Three Styles of The Scale

Scaling Categories	Items Number	Mean	Standard Deviation
Triple Scaling	20	0.2812	0.9964
Quatrain Scaling	20	0.4042	0.9509
Quinary Scaling	20	0.4542	0.9380

For revealing the difference of statistical significance between the mean values, several comparisons are drawn in table 13:

Table 13: Results of Several Frequent Comparisons Values of Study Sample Individuals' Performance on The Three Styles of The Scale

Scaling Categories	Comparisons	Mean Difference	Std. Error	Sig
Triple Scaling	Triple with Quatrain	* .1221	0.003042	.000
	Triple with Quinary	* .1721	0.003042	.000
Quatrain Scaling	Quatrain with Triple	* .1221	0.003042	.000
	Quatrain with Quinary	0.0500	0.003042	.106
Quinary Scaling	Quinary with Triple	* .1721	0.003042	.000
	Quinary with Quatrain	0.0500	0.003042	.106

* **Function at the significance level $\alpha=0.05$**

The table shows differences of statistical significance between the triple scale and the quinary in favor of the quinary scale and there are no differences of statistical significance between the quatrain scale and the quinary scale. The correlation coefficient has been calculated between the three styles of the scale as illustrated in table 14:

Table 14: Correlation Coefficient Between The Three Styles of The Sclae			
Correlation Coefficient	Triple Scaling	Quatrain Scaling	Quinary Scaling
Triple Scaling	1 352	0.537 * 352	*.513 352
Quatrain Scaling	** .537 352	1 352	*.560 352
Quinary Scaling	*.513 352	*.560 352	1 352

* Function at the significance level $\alpha=0.01$

T- Test sample has been used for the correlated samples and results are illustrated in table 15:

Table 15: Results of T- Test For The Correlated Samples		
Scaling Categories Comparisons	T	W
Triple with Quatrain	*5,33	0.62
Triple with Quinary	*7,24	0.52
Quatrain with Quinary	1,39	0.83

* Function at the significance level $\alpha=0.05$, Critical Value of 1.96 since $w = T =$

Table 15 shows differences of statistical significance between the triple scale and the quatrain scale in favor of the quatrain scale and between the triple scale and quinary in favor of the quinary scale, and there are no differences of statistical significance between the quatrain scale and the quinary scale.

4- Results and Discussion

4.1- Results discussion related to the first question: Does the assessment of reliability scale sign differ by the response categories number of the scales for each item?

Results show that the values of reliability coefficient Chronbach's Alpha increase by the increase of the scale categories number for the reliability coefficient values have increased in the light of the scale categories number from three categories to four, and from four to five categories. Moreover, there are differences of statistical significances among the value of these coefficients at the significance level ($\alpha=0.05$) since the calculated value M was (35..9444446) and it is statistically significant. These differences have been detected by using Z-Fisher test. Since the function between the quatrain and the triple scale is in favor of the quatrain, and between the triple scale and the quinary is in favor of the quinary because the reliability coefficient value of the quatrain scale is (.8246) is close to the reliability coefficient value of the quinary scale (.8506). Accordingly, this shows that the more the scale categories increase, the more the respondent has choices and thus he/she has to choose the state that suit him/her which leads to the reliability increase.

The previous results agree with a lot of previous studies such as, Al-Hazaimh's (1994) study, Al-Ghamidi's (2003), Master's (1973) study and Kim's (1998) study which show that reliability increase by the increase of categories number of the scale. While these results oppose Al-Akam's (1995) study and Matelle & Jacoby (1971) study. The researcher of this

study explains this result that when the categories number increase, the respondent has enough chance to determine what he feels. And the scale test expresses him accurately. This in turns increases the true variance value by calculating the error variance which leads to increase the reliability coefficient value opposing to the categories of less scale which doesn't express what does the respondent feel completely. The respondents choice of it because of having no other accurate alternative which leads to falling in errors. Thus, the error variance will increase at the expense of real variance leading to less reliability coefficient.

4.2- Results discussion related to the second question: Does the assessment of the validity scale sign differ by the response number of the scale?

The study findings related to the concurrent validity and how it is affected by the increase in the correlation coefficient value between the anxiety trait scale and the neurosis scale which represent the concurrent validity of the scale. The correlation coefficient value between the anxiety trait scale and the neurosis scale has reached (.65) on the triple scale and the correlation coefficient value between the anxiety trait and the neurotic trait has reached (.75) on the quatrain scale and between the anxiety trait scale and the neurotic scale has reached (.82) on the quinary scale. The statistically significant differences among the correlation coefficients between the values of the triple scale and the quinary scale are statistically different in favor of the quinary scale and not statistically significant between the triple scale values, the quatrain scale and the quinary scale. Regarding these results, they agree with Al-Ghamidi's (2003) and Mckelvie's (1978) study.

The results of this study emphasize that the correlation coefficient between the anxiety scale trait and the neurosis scale increases by the increase in the scale categories number. Besides, these results emphasize the importance of the tool and how it is suitable and accurate in measuring what needs to be measured. As regards, this supports the concurrent validity of the scale.

The increase in the scale categories number leads to an increase in the variance and correlation values which causes a constant increase in the value of the reliability and validity coefficients. This has been emphasized by Mckelvie's (1974) study since it shows that the small number of scale categories cause the lack of discriminated ability and validity of the scale.

In general, the findings of this study focus on the fact that the value of concurrent validity increases by the increase of the scale categories number used in the study.

4.3-Discussion of the third question findings

Does the assessment of the discriminated sign of each item in the scale differ by the response categories of the items?

Results pinpoint that the corrected correlation coefficient item of the triple scale is less than the corrected correlation coefficient of the quatrain scale and quinary scale. The corrected correlated coefficient of the quatrain scale is less than the corrected correlated coefficient item of the quinary scale. The single variance analysis has been used for the frequent measurements of the items' discriminated coefficient values since the findings illustrate that the items' discriminated coefficient values are affected by the categories scale increase from the triple scale to the quatrain scale and from the quatrain scale to the quinary scale. This explains that the corrected correlation coefficient value and the mean values increase by the increase in the categories number of the scale; whereas, the standard deviation values decrease by the scales' categories increase. Thses results show that the discriminated coefficient values of the items for the quinary scale is higher than the values of the discriminated coefficients of the items for other forms. This emphasizes that the scale with the most discrimination is the most reliable. By explaining the means, differences of statistical

significance are shown at the significance level ($\alpha=0.05$) among among the discriminated coefficients means of the item in the three styles of the scale categories for the style of the highest scale used in this study.

This study has revealed that reliability increases by the increase of scale categories and it will be statistically significant between the quatrain form and the triple form in favor of the quatrain form, between the quinary form and the triple form is in favor of the quinary form. It has shown that validity increases by the increase of the scale categories and it is statistically significant between the quinary form and the triple form in favor of the quinary. As a further sign of validity, the discriminated coefficients of the items have been calculated in each of the three forms. Accordingly, there has been statistical significance for the highest form. Moreover, the t-test of the correlated sample has revealed the existence of statistical significance between the triple scale and the quatrain scale in favor of the quatrain and between the triple scale and the quinary scale in favor of the quinary.

5- Conclusion:

In the light of the statistical analyses results, the researcher recommends the following: Paying attention to the anxiety trait scale characteristics and making sure that the suitable scale categories number gives the biggest values possible for both reliability and validity.

Conducting further studies when creating and decreasing the measurement tools by applying more styles of the scale differ in its categories to reach a suitable number.

- Experts in metrology should show more interest in the scale categories number.
- Using five scale categories of the anxiety trait.
- The necessity to determine the scale categories number scientifically and adequately.
- The increase in the response categories number of the personality traits scales.

References

- Abu Al-Sal, M. (2008). *Comparing between the verbal scale and the ordinal in Answering the self-report and assessment of others scales*. Unpublished MA Thesis. Jadara University, Irbid: Jordan.
- Al-'Akam, H. A. M. (1995). *The Effect of Categories Number Scaling of Likert's attitudes scale on psychometric characteristics*. Unpublished MA Thesis, University of Jordan.
- Abdel-Ghani J. M. (1996). Views of psychologists on fear and its stimulants. Cairo, Egypt: Al-Zahra' Library.
- Abdel-Rahman, S. (1998). *Psychological measurement*. Kuwait: Al-Falah Library.
- Al-Ghamidi, S. H. A. (2003). **The range difference of the measurement tool in the light of the change of the response alternatives number and teaching stage, studying the case of Likert's scale**. Unpublished M. A Thesis. Um Al-Qra University: Faculty of Education.
- Al-Sharifein, N. K. (1995). **The Effect of Distributing The Positive and Negative Items in The Attitude Scale on The Psychometric Characteristics And The Students Performance On It**. Unpublished MA Thesis, Yarmouk University, Irbid: Jordan
- Aiken , L. R. (1983). *Number of response categories and statistics on a teacher rating scale*. Educational and psychological Measurement, 43, 397-401.
- Aiken , L.R (1988) *Psychological testing and assessment* , Boston: USA: Allyn & Bacon .
- Anderson , D.& Brokewski , J.G (1987) *Experimental psychology*. Glenview Illinois : scott , Foresman and company.
- Chan, J. C. (1991). *Response order effects in liker- type scales*, Educational and psychological measurement, 151 (3), 531-540.

- Fishbein, M. and Ajzen, (1975). *Belief, attitude intention and behavior: An introduction to theory and research*. Reading MA: Addison-Wesley.
- Freeman, F.S (1962) *Theory and practice of psychological Testing*. New York, USA : Holting Rinehart and Winston .
- Hazaimah, A. (1994). *Factorial construction of Likert's attitude scale by the significance of scaling point's number*. Unpublished MA Thesis. Irbid: Yarmouk University.
- Jonson , W.L. and Dixon , P.N. (1984) *Response Alternative in Likert Scaling*. Educational and psychological Measurement, 44, 563-566.
- Kim , K. H. (1998) *Analysis of Optimal Number of Response*. Journal of Kams, 5-10.
- Lord , F.M.(1960). *Psychometric and Psychology*. Newly York: Academic press Inc .
- Matell , M.S. & Jacoby , J. (1971) . *Leshene an Optimal Number of alternatives for Likert Scales items*. Educational and Psychological Measurement, 31, 657-674
- Mckelvie, S. J. (1978). *Graphic rating scales- How many categories*. British journal of Psychology, 69, 185-202.
- Michael, A.(2006). *Psychological measurement II*. Damascus: University of Damascus Publications.
- Preston , C.C & Colman , A. M (2000) *Optimal Number of Response Categories in Rating Scale: Reliability, validity Discriminating Power and respondent Pre Rerences*. Acta Psychological, 104, 1-15.
- Spielberger, C.D. (1972). *Anxiety: current trends in theory and research*, Academic Press, New York.
- (Spielberger, C.D. and Diaz—Guerrero R.(Editors). (1982). *Cross-CulturalAnxiety*, Vol.2. Hemisphere, Washington,D.c.
- (Spielberger, C.D. and Sarason (1975) : *Stress and Anxiety*, vol. New York, John Wiley & Sons.
- (Spielberger, C.D. Gorsuch R . L . , Lushene R . , Vagg P.R . and Jacobs G.A. (1983) : *Manual for the State-Trait Anxiety Inventory (Form Y) Consulting Psychologists Press,Palo Alto*.