Use of Rasch-Andrich's Rating scale Model in the Calibrating of the Depression, Anxiety and Stress Scale (DASS-42) on high school students

Sonia AIOUADJ [1,*], **Rachid ZIAD** [2]

[1] Batna 1 University (Algeria), sonia.aiouadj@univ-batna.dz
[2] Tebessa University (Algeria), ziad.rachid@univ-tebessa.dz

Received:01/01/2023 **Accepted:**03/05/2023 **Published**: 28/07/2023

Abstract: The current study aims to calibrate items of the Depression, Anxiety and Stress Scale (DASS42) using Andrich's Rating Scale model emerged from the family of a the Rasch Unidimensional Measurement Model. The sample consisted of (404) high school students. Data were analyzed, using the statistical programs (Spss v21), and (winsteps v4.0.1). Parameters of the items, parameters of persons, and the standard error of measurement were estimated for both, in addition to the calculation of the statistical index of the Infit (MNSQ) and Outfit (MNSQ) of each item of the Andrich's Rating Scale model and all the 30 item lies After of the calibrated within an acceptable range of MNSQ value (0.84-1.17).

Keywords: Rasch-Andrich's Rating Scale model, Calibration, Depression, Anxiety, Stress, DASS42.

1- Introduction:

Anxiety and depression are often referred to as the common cold in mental disorders. Although this statement is accurate with regard to the prevalence of anxiety and depression, it tends to play down their consequences both personal and social. Among these consequences are the psychological, social and emotional sufferings of the person, as well as the imbalance of relationships with others, as well as the economic burdens on society. In view of the serious and pervasive effects of anxiety, and depression, the development of frameworks and models for understanding and treating it is extremely important. This is why empirical knowledge has recently increased; psychologists and psychiatrists have developed and validated models of anxiety and depression for many theoretical approaches (David et al, 2015, 15)

This was also confirmed by Dominique (2012) who sees that psychological stress and anxiety today are among the most common problems mental health is prevalent, and that stress and stressful events for a long time lead to the emergence of disorders anxiety, and he adds that all studies indicate the presence of anxiety and depression as a common and high disorder, as can juveniles stressors can have a significant impact on the development of an anxiety disorder, whether in childhood, adolescence, or at a later time (Dominique, 2012, 1-75).

Ali (2008) adds that depression, anxiety and stress psychological disorder is one of the most important psychological disorders most commonly treated by psychologists at the present time (Ali et al 2008). However, what is most striking is the relationship between anxiety and depression itself, as there are many commonalities between these two disorders, and this mutual phenomenon raises doubts about whether the diagnosis is a concern or depression. In fact, there is an empirical distinction, and at the level of symptoms, the correlation coefficient between anxiety and depression tends to have a value of 0.61 on the self-report scales (1985, Dobson), while the average The association rate between major

* Corresponding author

depressive disorder and various anxiety disorders is more than 50% in samples of the general population and clinical samples (David et al, 2015, 53). The results of more recent studies also confirm that psychological stress can be generated by the emergence of anxiety and depression. This prompted specialists to pay more attention to the sensitivity perspective to be affected by psychological stress as a conceptual framework for understanding the pathological causes of the depression, to a lesser extent, understanding the causes of anxiety (David et al, 2015, 140), while many studies have found that anxiety precedes depression (Kendall and Brady, 1992), (Dobson, 1985)

Depression is one of the most common mental disorders among adolescents, with estimated lifetime prevalence between 15% and 20%. Depression also appears to be in common with other psychological disorders. And after studying six community studies Castello and Angold (1993) concluded that the presence of depression in adolescence increases the risk of the probability of having another disturbance at least twenty times. Anxiety is the disorder most commonly associated with depression, and anxiety problems precede depressive disorder (2002, Essau) (David et al, 2015, 283)

Meanwhile, several studies have linked negative life events and chronic stress with simultaneous occurrence of depression as in the study of (Essau, 2002; Lewinsohn, Clark, Seeley, and Rohde, 1994) and also was confirmed by the British Peoples' Commonwealth of Nations (Commonwealth) survey of adolescent health in 1997, (Schraedley 1999, Hayward and Gotlib) that the presence of many triple stressful life events in those diagnosed with high levels of depressive symptoms, the occurrence of depression in the majority of affected people is associated with high levels of depression; major depressive disorder is associated with severe stressful life events, and environmental stresses are associated with the onset of depressive symptoms in childhood and adolescence. (David et al, 2015, 410-289)

Consistent and reliable measurement tools are the cornerstone for researchers and clinical professionals who care about research and practice activities for the prevention of mental disorders, which are related to psychiatry and clinical psychology, where these scales emerge from ingenious theoretical structures that are the basis for answering research questions in a manner where the scales are internally consistent or homogeneous, their use can be generalized in the future. (David et al, 2015)

Among the most important scales that are commonly used in diagnosing the structure of negative emotions are depression and anxiety scale And Stress (DASS) (1995, Lovibond and Lovibond), which is available in two forms, one of which consists of 42 items and the other is short consisting of 21 items, suitable for estimating symptoms of depression and anxiety, as well as the dimensions. It also allows obtaining the comorbidity rate for both anxiety and depression, So, It is considered a sufficient measure to estimate and distinguish at least two types of symptoms, as it has proven its worth through the number and prevalence of its use in many international studies that have been reviewed, such as the study of (Lovibond.F and Miriam), and (Brown et al, 1997) and (webster.S et al, 2013) and his achievement of good psychometric indicators in different environments, In addition, it is a low-cost and time-consuming scale that measures three disturbances simultaneously, in addition to its use as an external test for measuring collateral validity in many published studies and researches.

In our current study, we relied on the use of the Andrich scale model in calibrate the scale, which is one of the standard models and recent psychometric developments that have received remarkable attention from psychometric and educational scientists recently, which are closely related to the so-called latent trait models of measurement, from which the theory has emerged the Test item response, and one-parameter logistic Rasch model, Which takes the form of multiple responses Likert type (ploytomous) with increments separated by equal distances, and this model identifies a group of the items shares the structure of the rating scale, whereby the same response alternatives are selected for all items, compared to partial grading model that determines for each item its own grade scale (parzer & Fischer, 1991). The grading scale is characterized by the presence of thresholds that express the boundaries between the steps, and they are fixed across the item, and the model is based on the idea that

each of the items of the scale carries a total emotional charge, and the model estimates this charge for each item is according to the probabilistic mathematical function adopted by the model (De Ayala, 2009)

Despite the importance of this statistical model, studies in the Arab environment in general and in Algeria in particular which I dealt with is rare - within the limits of the researchers' knowledge - as the process of codifying scales using the scale model, which it is an evolution of the Rasch model in dealing with multiple scales in which it is not limited to binary response as in the Rasch model, so the scale model is the most appropriate model for the response to the test item to analyze and grade psychological and educational scales at the present time.

Among the most important studies that focused on and employed models of calibrate scales emanating from the Rasch model in calibrate Depression, Anxiety and Stress Scale: Study of both (2009, al et Tracey) which used the Scale Negative Emotions Depression, Anxiety and Stress (DASS21), which was applied to a sample of 420 people in Australia, using the Rasch model credit partial, and the program (RUMM2020). The results showed the deletion of three items from the scale, which did not support the unidimensionality of the total scale, but to two scales (anxiety and depression). And the study of each of (Al et Oleg, 2018), in which Depression, Anxiety and Stress scales were used (DASS21) the shorten copy, and the Partial Grade Grading Form, and the program (RUMM2030) for (Andrich et al, 2009), on a sample of (400) New Zealand university students, showed the results are multiple dimensions of the scale with the deletion of the word number 05 from the scale. And the study of both (Parkitny and Luke 2011) using the Depression, Anxiety and Stress Scale (DASS21) (Abridged Version, and Rating Model) Partial Undergraduate Rasch-Score, on a sample of (154) Australian volunteers with chronic pain for the beck, the results indicated that they do not support the unidimensionality of the overall scale, opposite to the three subscales (psychological stress; anxiety; depression), the results of which were good. It is evident by reviewing some studies that dealt with evaluating competency the psychometric scale for depression, anxiety, and psychological stress using the Andrich-Rasch calibrate scale model was used The short version of the scale, and we did not find, within the limits of the researchers' knowledge, a study that used this methodology in calibrate Depression, Anxiety and Stress Scale full version, and the results of these studies were conflicting in the process of calibrate the scale, which varied according to the samples and the environments in which each study was conducted.

1.1- The study problem and its questions:

The problem of the current study is the need to develop an accurate measurement tool that diagnoses negative emotional states. Depression and anxiety are of great theoretical and clinical importance. Conceptually, depression and anxiety are two disorders completely balanced, but the clinical interaction between the two conditions has been confirmed by many researchers and practitioners clinicians. The concept of psychological stress also posed additional problems in the study of negative emotional cases as a pattern emotional reaction or condition, the psychological stress has a clear affinity with anxiety, and the two concepts usually do not differ between them Cleary. The Depression, Anxiety and Stress Scale have gained worldwide fame in measuring and diagnosing emotional disorders Negative, a few studies have been conducted - according to the researchers' knowledge - in the Arab environment to verify its psychometric properties What was presented was according to the classical theory of measurement, which proved to be insufficient in achieving objectivity in measurement, through the dependence of its measurement on the characteristics of the test used (Swaminathan and Hambleton, 1985).

Although the singular response theory and its models, including the Andrich-Rasch calibrate scale model, which is one of the Unidimensionality models of the Rasch model, It is dedicated to Likert data, and the objective advantages of this model in building and developing such a tool The measurement used in the current study, however, the researchers

noticed a lack of theoretical literature in Arabic studies in the use of this statistical model in calibrate the depression, anxiety and psychological stress scale (DASS42 .).

It is possible to define the problem in the form of questions that the study aims to answer:

1. What is the Calibrating of the items difficulty of the Depression, Anxiety and Stress Scale (DASS42) using Rasch-Andrich Rating Scale model?

2. What are the validity and reliability indications available for the Depression, Anxiety and Stress Scale items (DASS42) after calibrate its items using Rasch-Andrich Rating Scale model?

2 - Objectives of the study:

1. Scaling the Depression, Anxiety and Stress Scale Item (DASS42), using the Andrich Scale of Estimation model, which is derived from the Unidimensionality logistic Rasch model.
2. Verification of the psychometric efficiency of the scale through the validity and reliability indicators of the Depression, Anxiety and Stress Scale (DASS42).

3 - The importance of the study:

The study gains its importance in that it provides the field of diagnosing negative emotional disorders, depression, anxiety and psychological stress among secondary school students in the Arab environment in general and the Algerian environment in particular, with an adapted and developed tool that has objective psychometric properties, through applying the Rasch-Andrich's Rating Scale model in the calibrating, which stems from the Unidimensionality logistic Rasch model, which helps reduce Diagnostic and classification errors.

4 - Defining the study terms:

4-1- Item Calibration:

The calibrate process is concerned with determining the correspondence between the observed data and the locations of people on the latent variable. If we locate persons on the latent variable, we can then compare them to each other (De Ayala, 2017, p.03).

4-2- Andrich's Rating Scale Model (RSM):

The Grade Scale Model (RSM) is considered by Andrich (1978). A model derived from the general model of the Rasch model (1961), Andersen (1977), as well as (Wright and Masters, 1982). It is a linear model describing the probability of an individual responding (n), on a given item according to a scaled Likert scale (i), and categories of graded responses (x) (Andrich, 1978). The mathematical model for this probability, as illustrated by the following equation, includes three parameters: individual ability (β_n) singular difficulty (δ_i), and step difficulty ($\boldsymbol{\tau_j}$) (that is, the threshold between two successive categories, x and x-I). Calibrating the scale data according to this model produces an estimate for each separate parameter, as well as a standard error for this estimate for each person, for each item, and the scale of the step in the context of this measurement, which is given by the following equation:

$$P(X_{ni}=x)=\frac{\exp\sum_{j=0}^{x}[\beta_n-(\delta_i+\tau_j)]}{\sum_{k=0}^{m}\exp\sum_{j=0}^{k}[\beta_n-(\delta_i+\tau_j)]},x=0,1,..m$$

Where: P ($\mathbf{X_{ni}} = \mathbf{x}$) is the probability of an persons responding (n), on an item (i), in categories of graded responses (x), and a threshold ($\boldsymbol{\tau_j}$), and each item has sections of graded responses (**m + 1**). (Wright & Masters, 1982, 50)

5- Method and Tools:

5-1- Study Methodology:

In analyzing the study data, the researchers relied on the descriptive analytical method, the Rasch model, and the Andrich Rating Scale (RSM) in order to rank the Depression, Anxiety and Stress Scale (DASS42).

5-2- Study sample:

An available sample (Convenience Sample) was selected from secondary school students in El-Oued State, during the first semester of the academic year 2017-2018. It consists of (404) male and female students distributed among the first, second and third-year students of all secondary schools, their mean age is 18 years, and a standard deviation of 2.16, of whom (266) are female students, or 65.8%, and (138) male students are (34.1%) out of the total sample.

5-3- Study Tool:

In the current study, we used the data of the Depression, Anxiety and Stress Scale (DASS42) for (Lovibond, Lovibond; 1995). The scale, in its original form, consists of 42 items. It is a self-report response scale designed to measure the main symptoms of negative emotional reactions, psychological stress; anxiety; Depression. The items on the scale are divided into three sub-scales: psychological stress with 14 items; anxiety in 14 items; depression with 14 items; The answer requires choosing one of four alternatives, namely: (0) It does not apply to me at all, (1) It applies to me some or a little of the time, (2) It applies to me significantly or some of the time, (3) It applies to me very often or most of the time, you rate the applicability of this sentiment in the past week.

6 - Results and Discussion:

Before carrying out the process of calibrating the scale, the researchers verified the conditions and assumptions of the Rasch model, the results of which were as follows:

6 -1 – Unidimensionality:

This assumption is one of the basic assumptions on which the rating scale model is based. The model assumes the presence of one feature that explains the persons performance on the negative emotional scale (psychological stress, anxiety, depression). For this, the results of the analysis of the Rasch model were used to analyze the main components based on the Rasch Principal Components Analysis of Items Residuals (Rasch PCAR), which shows the differences between the dimensions, which is shown in table 1.

Table.1. factor analysis of the main components of the residuals using the "Rasch" model.

The magnitude of the variance of standard residuals in Eigen units	potential root	The observed	The expected
Overall variance in responses.	64.6119	%100	%100
The variance explained by the principal factor (Rasch model estimates).	22.6119	%35.0	%35.4
The variance explained by persons.	5.8382	%9.0	%9.1
The variance explained by items.	16.7737	%26.0	%26.3
The unexplained variance.	42.0000	%65.0	%64.6
The variance explained by the second factor (the first in residuals).	2.9405	%4.6	

Source: winsteps v4.0.1 . output

Based on the results of Table 1 and to judge the Unidimensionality of the factor analysis of the main components of the residuals using the "Rasch" model, we relied on the following criteria:

1. The first test, as a general rule, is if the value of the variance explained by measures is greater than or equal to (60%) or (50%), as stated in the program guide, and guided by what was specified by the developer of the (winsteps) Linacre Mike in a direct contact with him, as he determined the range from 20% to 80% as an acceptable range and a strong indicator of Unidimensionality, which was achieved in the current study, where we recorded the value (35%) for this test.

2. The second test is that the percentage of the variance explained by the second factor (Unexplained Variance in 1st contrast) should be less than (5%), and this was achieved in the current study with the value (4.6%). This is another evidence of Unidimensionality.
3. The third criterion is the value of the potential root of the contrast ratio, which is explained by the second factor (Unexplained Variance in 1st contrast), as it must not exceed or be less than (3), and this was also achieved by recording the value (2.9405) which is less than (3). (Linacre, 2006)

The efficiency of the scale categories or response sections of the depression, anxiety and psychological stress scale was also evaluated: (not applicable - slightly applicable - some time - applicable a lot), where the percentage of the category's appearance in the observed data and the mean estimates of persons within the category and the Rasch-Andrich parameter were calculated, as Table 2 shows .

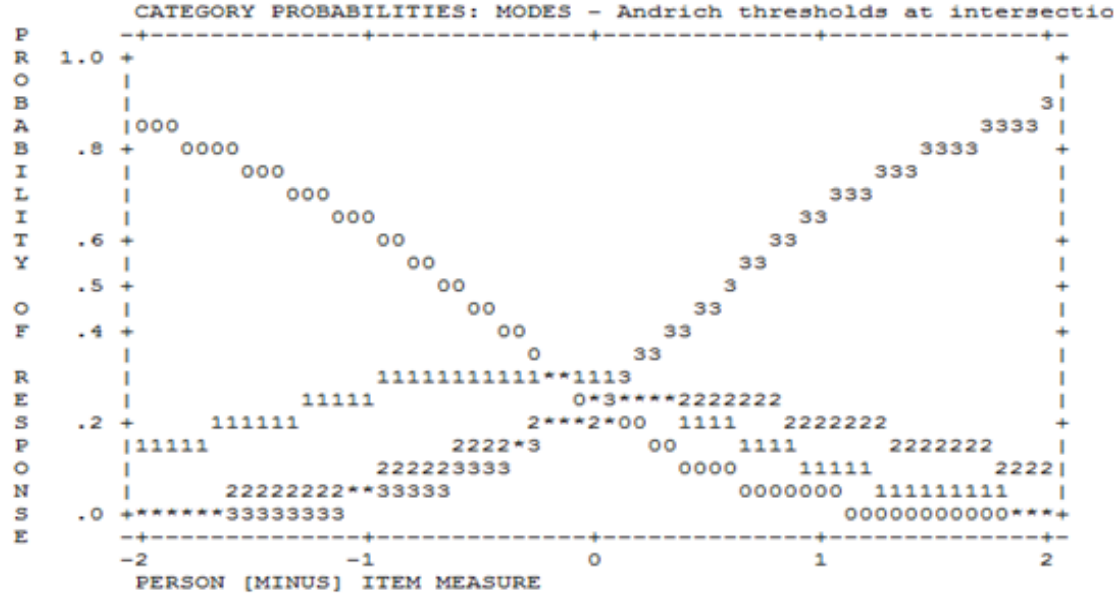
Table.2. Percentage appearance of the category in the observed data and mean estimates of persons within the Rasch-Andrich category

Response categories.	The assigned grade	The observed percentage	Average capacity		Convergent and divergent fit statistics for categories		Andrich-thresholds	Measurement categories
			Observed	expected	*INFIT-MNSQ*	*OUTFIT-MNSQ*		
1	0	48	-1.11	-1.11	1.02	1.07	-	-1.63
2	1	25	-0.57	-0.58	0.99	0.92	-0.16	-0.40
3	2	13	-0.20	-0.18	0.99	0.95	0.30	0.45
4	3	15	0.25	0.25	1.00	1.03	-0.14	1.57

Source: winsteps v4.0.1 . output

It is clear from Table 2 that the mean ratings of the trait in the four categories are arranged ascendingly according to the degree of the category, as expected, and that all response categories are statistically appropriate according to the convergent and divergent appropriateness scales. We note that the increase monotonically for the calibrate of the Rasch-Andrich thresholds was not increasing with the increase in the category, which means that the assumption of the scale model is not fulfilled. The following figure 1 illustrates this:

Fig.1. shows the response probabilities curves for the four categories



Source: winsteps v4.0.1 . output

Figure 1 indicates that the distinction between category 1 "applies to a little or some of the times" and category 2 "applies to a significant degree or sometimes" was not clear to the study sample. This suggests that these two categories should be combined into one category, which is: "Applies to some times." The following table 3 shows the results of the modification after deleting one of the four response categories.

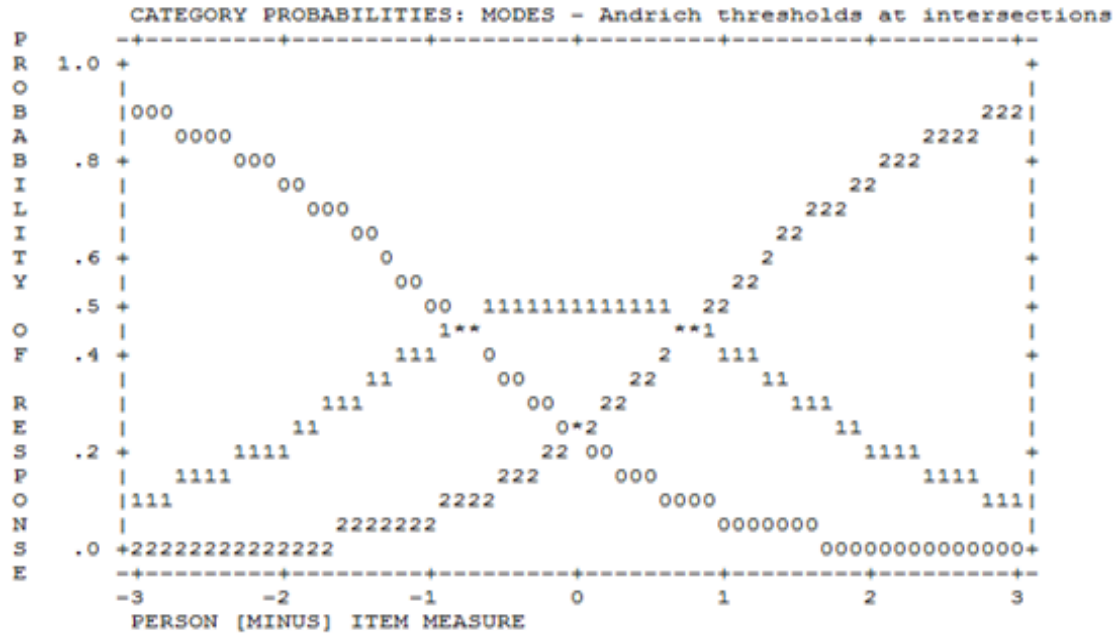
Table.3. Percentage appearance of the category in the observed data and average estimates of persons's within the category by Andrich

Response categories	The assigned grade	The observed percentage	Average capacity	Convergent and divergent fit statistics for categories	Andrich-thresholds	Measurement categories		
			Observed	expected	*INFIT-MNSQ*	*OUTFIT-MNSQ*		
1	0	48	-1.53	-1.53	1.03	1.05	-	-2.01
2	1	37	-0.60	-0.59	0.98	0.94	-0.78	0.00
3	2	15	0.38	0.36	0.98	0.99	0.78	2.01

Source: winsteps v4.0.1 . output

It is clear from Table 3 that the mean estimates of the trait in the three categories are arranged ascendingly according to the degree of the category, as expected, and that all response categories are statistically appropriate according to the convergent and divergent appropriateness scales, as the MNSQ statistics for the appropriateness of the categories did not exceed the appropriateness limits (0.70-1.30), and We notice that the increase monotonically for the Calibrate of the Rasch-Andrich thresholds increases with the increase of the category, which means that the assumption of the scale model is fulfilled. And each of the three categories responded to more than 10 people recommended by it (Linacre, 2002), and the following figure 2 illustrates this.

Fig.2. shows the response probabilities curves for the three categories.



Source: winsteps v4.0.1 output.

Figure 2 shows the probability curves for the three categories, as it is clear that the thresholds are no longer overlapping, and that the values in each category reached a peak and within a satisfactory range. Category (0) is most likely to have ratings (negative emotional trait) of less than -0.78; Category 1 is for estimates between (0.78) and (0.78), that is, the sum of the step difficulties (1.56 logit), a value greater than the minimum value (1.4 logit), and less than (5.0 logit) for calibrating the difficulty of the steps recommended by Linacre. (Linacre, 1999) for three category, the lowest value for five categories (1.0 logite); Finally, Category 2 was most likely for people with scores above 0.78. Therefore, all other criteria have been met to diagnose the optimal number of categories.

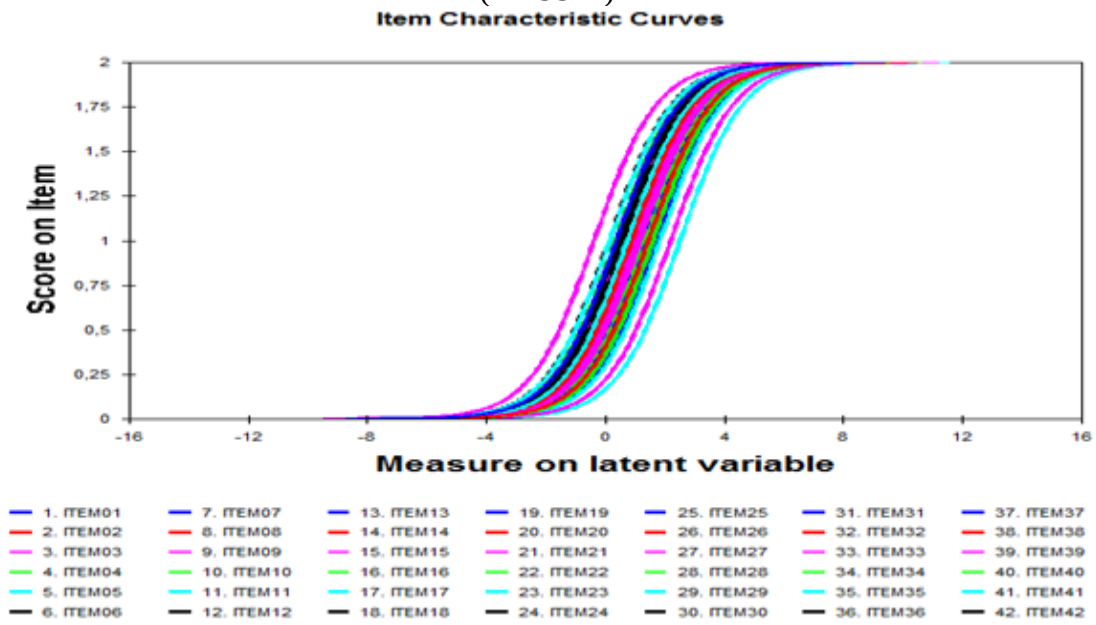
6-2- Checking the Local Independence Assumption:

Hambelton and Swuminathan (1985) see that this assumption is equivalent to the Unidimensional assumption, which means that the assumption is valid.

6-3- Assuming Equal Discrimination Coefficients:

Discrimination coefficients were calculated for the items of the scale by calculating the serial binary correlation coefficient (r_{pb}) between the performance of persons's on the item and their performance on the scale as a whole, which is considered an indicator of the degree of discrimination of the items. Hambelton and Swuminathan (1985) indicated that in order to achieve assuming the equivalence of the discrimination indicators and their conformity with the model, their value must be realistic within the limits of the range (mean discrimination coefficients = 0.15), and since the mean discrimination coefficients are (0.52), the range limits range between (0.38) and (0.66), and therefore the values of the sequential binary correlation coefficients All were within the perfect range. The researchers also used the (winsteps) program to extract the item characteristics curve, which is illustrated in the following fig3.

Fig.3. shows the item characteristics curve of the Depression, Anxiety and Stress Scale (DASS42)



Source: winsteps v4.0.1 output.

It is clear from Fig.3 that the Slope is equal, as Linacre (2012) sees that the mean values of item discrimination in the "Rasch" model should be close to (1.0). This is what was achieved in our current study, as the mean discrimination values are equal to (0.98), which is very close to one. (Linacre, 2012)

6-4- Guessing:

Regarding the hypothesis of the low guess index, we can observe the drawing of the item characteristics curve (the non-linear regression lines of the degrees on the scale), that the lower asymptote approaches zero in all the item and this happens when the persons does not resort to the answer at random, which is an indicator of good matching the item of the model, and it cannot be said that the speed factor played a role in answering the item, when enough time was left for the respondent.

- Presentation and discussion of the results of the first question:

What is the Calibrating of the items difficulty of the Depression, Anxiety and Stress Scale (DASS42) using Rasch-Andrich Rating Scale model?

Data on the responses of the study sample members (404) of secondary school students, and (42) item of the total scale were entered on the winsteps program, and the results were extracted according to the following steps:

1. Elimination of persons who do not fit the model, according to the following criteria:

-Excluding persons whose fitness statistics value is less than (-2), as this means that the answers of these persons are similar, which indicates the inaccuracy of the responses.

-Excluding persons whose fitness statistics exceed (+2), as this means that the answers of these persons are similar, which indicates the insincerity of the responses, as they chose an alternative that expresses the negative emotion trait (more or less) than their trait level, and

this may be due to neglect and their lack of seriousness in responding (Wright & Linacre, 1998). This step resulted in the deletion of (71) persons.

2. The analysis was re-analyzed again to delete the inappropriate terms for the model, according to the following criteria:

-Deleting items for which the appropriateness statistics values are less than the value (-2) as this means that those items are not independent from the rest of the scale items, or that they measure another variable very similar to the one supposed to be measured.

-Deleting items whose values of relevance statistics exceed the value (+2), as this means that there is a defect in the construction of the items, or that it measures another variable. This resulted in the deletion of (12) items and Table 4 shows that.

Table.4. items difficulty and standard errors in descending order according to items difficult.

Item	Raw score	Point bisreal PTBIS-CORR	Difficulty	SE of model	Infit		Outfit	
					MNSQ	**ZSTD**	**MNSQ**	**ZSTD**
23	89	0.42	1.56	0.13	1.12	1.2	1.08	0.5
15	116	0.46	1.18	0.11	0.93	0.8-	0.94	0.4-
17	144	0.48	0.85		1.14	1.6	1.08	0.7
36	145	0.51	0.83	0.10	1.00	0.1	0.98	0.2-
19	155	0.38	0.73	0.10	1.28	3.3	1.31	2.7
28	165	0.57	0.62	0.10	0.99	0.2-	0.91	0.9-
04	169	0.47	0.11	0.10	1.14	1.7	1.14	1.4
34	174	0.51	0.53	0.10	1.05	0.7	1.13	1.3
02	184	0.39	0.44	0.10	1.11	1.04	1.16	1.7
41	184	0.50	0.44	0.10	1.13	1.07	1.14	1.4
38	185	0.54	0.43	0.10	1.14	1.08	1.06	0.6
07	186	0.48	0.42	0.10	1.03	0.4	1.02	0.3
20	191	0.53	0.37	0.10	1.01	0.2	0.98	0.2-
31	200	0.57	0.29	0.10	0.95	0.7-	0.90	1.1-
10	201	0.46	0.28	0.10	1.17	2.2	1.23	2.4
03	206	0.46	0.24	0.09	1.01	0.1	1.12	1.3
37	206	0.57	0.24	0.09	1.06	0.9	1.00	0.0
25	212	0.58	0.18	0.09	1.03	0.4	0.94	0.6-
21	222	0.54	0.10	0.09	1.14	1.9	1.17	1.9
33	224	0.61	0.08	0.09	0.92	1.2-	0.88	1.5-
24	227	0.59	0.05	0.09	0.86	2.0-	0.83	2.2-
39	242	0.64	0.07-	0.09	0.74	4.1-	0.71	3.9-
08	248	0.51	0.12-	0.09	1.05	0.8	1.07	0.9
26	252	0.64	0.16-	0.09	0.80	3.1-	0.80	2.7-
16	254	0.55	0.17-	0.09	0.97	0.4-	0.97	0.4-
22	257	0.60	0.20-	0.09	0.87	1.9-	0.87	1.7-
30	259	0.47	0.21-	0.09	1.02	0.3	1.04	0.5
05	272	0.50	0.32-	0.09	0.97	0.4-	1.13	1.7
18	278	0.52	0.37-	0.09	1.16	2.3	1.11	1.4
32	281	0.46	0.39-	0.09	1.09	1.4	1.10	1.3
27	282	0.63	0.40-	0.09	0.71	4.8-	0.76	3.4-
35	299	0.48	0.53-	0.09	1.17	2.5	1.19	2.4
14	303	0.44	0.56-	0.09	1.22	3.2	1.21	2.6
06	305	0.47	0.58-	0.09	0.99	0.2-	0.99	0.2-
13	305	0.66	0.58-	0.09	0.73	4.5-	0.73	4.0-
29	309	0.51	0.61-	0.09	0.94	0.9-	0.92	1.0-
01	311	0.49	0.62-	0.09	0.94	1.0-	1.01	0.1
40	318	0.62	0.68-	0.09	0.89	1.7-	0.85	2.1-
42	319	0.55	0.69-	0.09	0.79	3.4-	0.80	2.9-
11	340	0.63	0.85-	0.09	0.83	2.7-	0.80	2.7-
12	343	0.59	0.87-	0.09	0.94	0.9-	0.90	1.3-
09	417	0.45	1.46-	0.09	1.14	2.0	1.20	2.2
M	**237.6**	**0.52**	**0.00**	**0.09**	**1.00**	**0.1-**	**1.00**	**0.1-**
SD	**68.5**	**0.07**	**0.60**	**0.01**	**0.14**	**2.0**	**0.15**	**1.8**

Source: winsteps v4.0.1 output.

After excluding the non-conforming persons, the non-matching items from the Depression Scale 10; 24; 26; 13; 42, non-conforming Anxiety Scale 19; 9, non-conforming

Stress Scale 39; 27; 35; 14; 11, and for the purpose of confirming the conformity of the data resulting from the responses of the sample members to the scale and verifying the objectivity of the scale in its final form (30 items), the analysis was re-analyzed to obtain estimates that were free of the items difficulty and the abilities of the persons, and Table 5 shows the results of the analysis for the values freed from the abilities of the persons.

Table.5. shows the results of the analysis of the values freed from the abilities of persons.

Statistics	Raw score	Ability	Standard Error of the model	Infit		Outfit	
				MNSQ	**ZSTD**	**MNSQ**	**ZSTD**
Mean	19.9	0.93-	0.34	1.00	0.0	1.00	0.0
SD	3.4	1.13	0.11	0.28	1.1	0.28	1.1
High score	58.0	3.63	1.01	2.03	3.6	1.95	3.4
Low score	1.0	4.34-	0.27	0.39	3.8-	0.42	3.5-

Source: winsteps v4.0.1 output

It is clear from Table 5 that the final estimates liberated from the abilities of persons ranged between (1-58), the mean power distribution (-0.93) logit, and the standard deviation (1.13) logit, and the standard error of the mean of the power estimates was (0.11) which is close to the ideal situation that the model assumes, which indicates the accuracy of locating persons on the trait continuum, and Table 6 shows this:

Table.6. shows the results of the item-free analysis

Statistics	Raw score	difficulty	Standard Error of the model	Infit		Outfit	
				MNSQ	**ZSTD**	**MNSQ**	**ZSTD**
Mean	221.2	0.00	0.10	1.00	0.1	1.00	0.0
SD	62.4	0.56	0.01	0.08	1.1	0.10	1.1
High score	343.0	1.40	0.12	1.15	2.1	1.17	1.9
Low score	89.0	1.00-	0.09	0.87	1.9-	0.84	2.2-

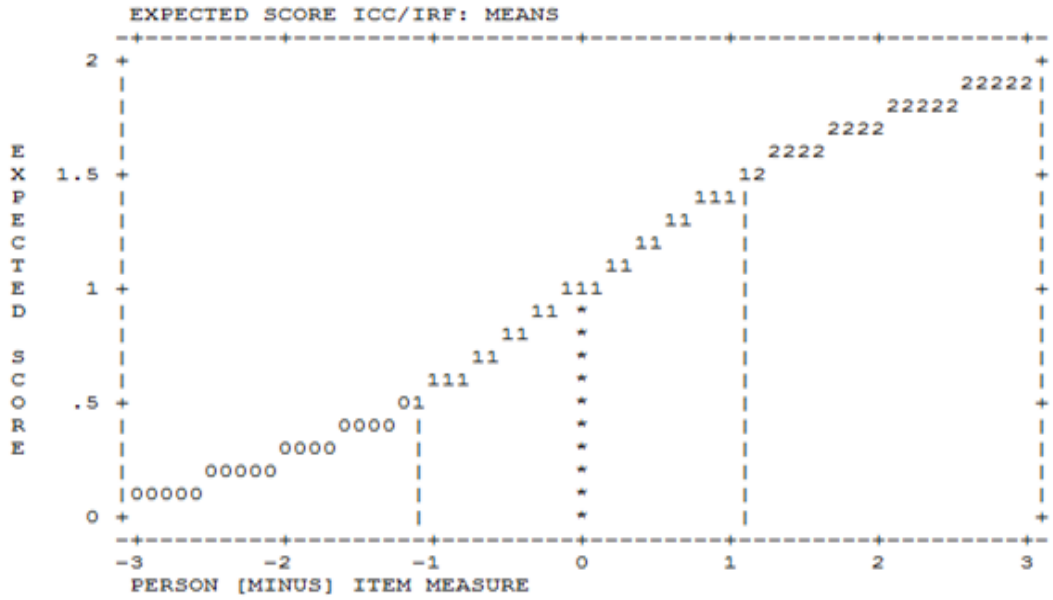
Source: winsteps v4.0.1 . output

Table 6 indicates that the mean power distribution amounted to (0.0) logit, standard deviation (0.56) logit, and the values of the estimates freed from items difficulty ranged between (-1.00) to (1.40), and the standard error of the mean for estimates of difficulty was (0.10), which is a low value , which indicates the accuracy of the items difficulty estimates, and the values of the 30-item scale items parameters were estimated, using the Unconditional maximum likelihood estimation (UCON) method, for a more accurate estimate of abilities and difficulty, and to reduce the error in estimating the difficulty of items. The values of the statistical fit indicators for the mean of divergent and convergent squares (MNSQ) for the items of the scale indicate that they are close to the ideal situation assumed by the model, which is the value (1), where its values ranged between (0.84-1.17) with an mean (1.00) and a standard deviation (0.10). , which is in the field of productive measurement defined by (Wright and Linacre, 1994) with values (0.70-1.30) or values (0.50-1.5) in general and close to one, and it is a statistical indicator that is more sensitive to unexpected behaviors from the model (extreme such as guesswork and errors unexpected) that affects responses to item that are far from the persons ability level. It is also considered another evidence of the one-dimensionality of the negative emotional reactions scale. Therefore, a value greater than (1.0) for Outfit-Underfit or MNSQ < 0.7/-2 means that there is a large discrepancy or overlap in the data that is not justified, and a value less than (1.0) for (Infit-Overfit) or (MNSQ) indicator (MNSQ >1.3/2 means that the model is over-data and may inflate the stability statistics (William j et al, 2014).

Items polarity also tests the extent to which the items works in one direction, to determine the measured hypothesis. The result of the analysis showed that the point measure correlation PT-MEA CORR is positive. Its values ranged between 0.38 for item 19, and the value 0.66 for item 13, with an mean of 0.52 and a standard deviation of 0.07, which indicates that these items measure the hypothetical formation that he intends to measure, is negative emotional reactions (psychological stress, anxiety, and depression).

The abilities and Response Possibilities

Fig.4. shows the Expected Item Characteristic curve (ICC) or the IRF by the three response divisions (categories)



Source: winsteps v4.0.1 output.

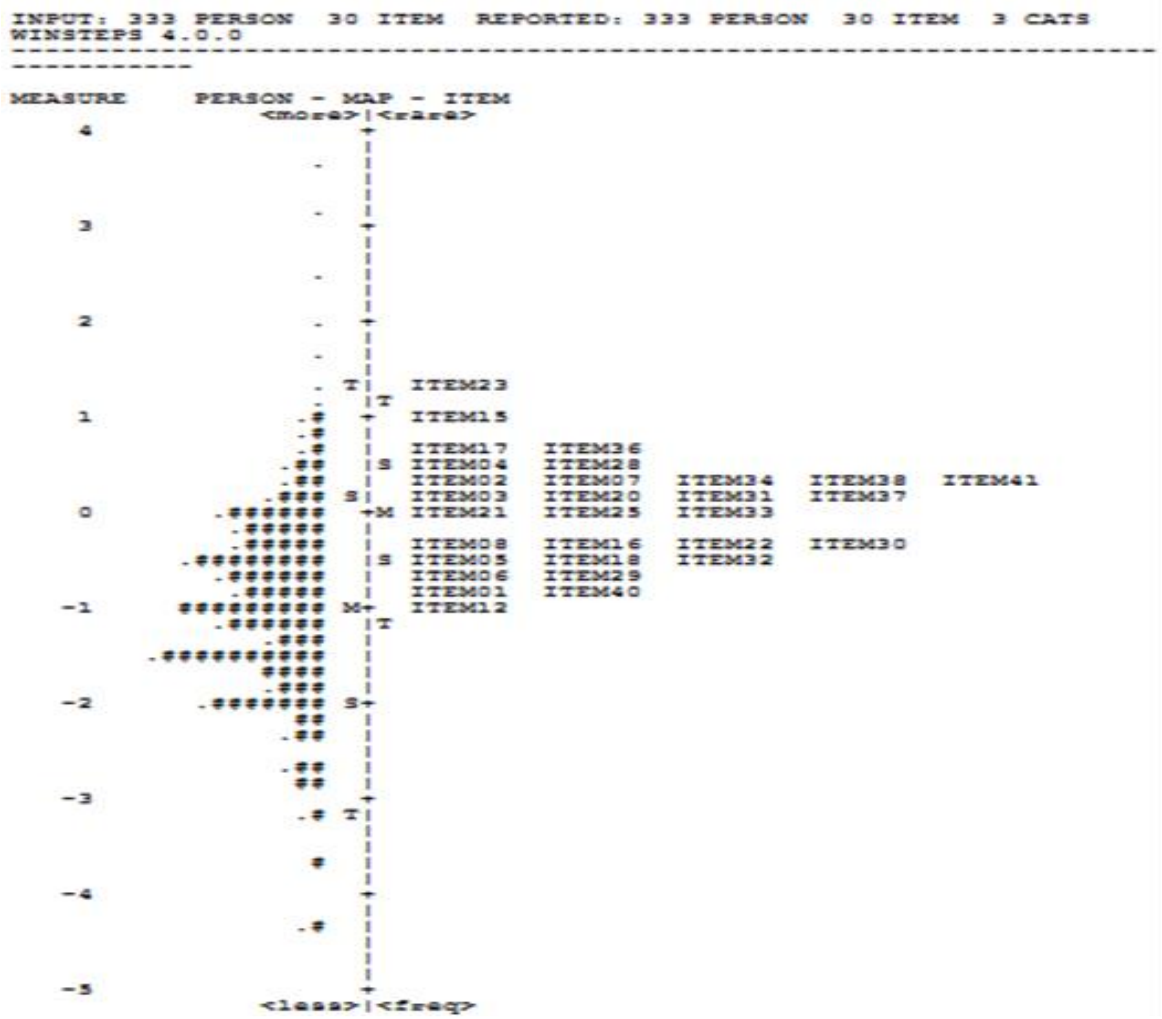
It can be seen from Figure 4 of the ICC (expected scores of the model), which is also called the IRF, which indicates the Rasch model's expectation of performance on the overall scale, and for the difficulty of the items, it turns out that it gradually increases according to the levels of ability, which it determines Dashed lines for Thresholds, the middle of which (average difficulty) on the ability scale corresponds to the expected values of the model (0.5), and the probability of a correct answer is low at lower levels of ability, and approaches the correct one at higher levels of ability. Therefore, the difficulty is a guide to the position of the persons (Location index) on the characteristic of its curve. The second characteristic is discrimination, which appears or is reflected in the steepness of the characteristic curve of the persons in the middle. The steeper the curve, the more able the persons to discriminate, and the more flatter the curve, the less ability of the persons to distinguish because the probability of response at low levels of ability approaches it in the case of high levels. (Beeker, 2010)

-Presentation and discussion of the results of the second question:

What are the indications of validity and reliability available for the items of the Depression, Anxiety and Stress Scale (DASS42) after calibrate its items using the Andrich rating scale model, which is derived from the Rasch model?

To verify the structural validity of the scale, the researchers used the Wright map after the final calibrate of the scale, and Figure 5 shows the results of that.

Fig.5. shows a Wright map



Source: winsteps v4.0.1 output.

Wright's map in Figure 5 clearly indicates that there is a perfect fit between the distribution of the estimation of depression, anxiety and psychological stress scale items from the right side of the map, and the calibrate of persons from the left side of the map. Although the calibrate of the items difficulty is clearly from the range -1.00 and 1.40 logit, it may appear that some spaces or gaps at some other levels of the variable shown by the Wright map, which can be verified through the relationship between the space distance between the difficulty of the two items and the total error The standard error of the two, for which the difference between the difficulty of the two terms specified for the space must be less than the sum of the standard error of measurement for both and the spaces are:

The first space: the difference between the two items difficulties 21 and 8 is 0.03 and the sum of the standard error of the measurement is 0.22, that is, the standard error of the measurement covers the space.

The second space: the difference between the difficulty of the words 23 and 15 is 0.38, and the sum of the standard error of the measurement is 0.24, meaning that the standard error of the measurement does not cover the space.

The third space: the difference between the difficulty of the words 15 and 17 is 0.33 and the sum of the standard error of the measurement is 0.22, that is, the standard error of the measurement does not cover the space.

Thus, the standard error of measurement for the difficulty of each of the two items constituents of the space covers the distance of the first space between them, but it does not cover the distance of the remaining second and third spaces with a small difference. Which represents the extent of this connection, which are negative emotions (depression, anxiety, psychological stress). This indicates the quality of the scale item calibrate.

Scale Reliability: Two types of coefficients were obtained: Person Reliability and Item Reliability. Reliability according to the response theory to the items means accuracy in estimating the location of each of the persons and item on the attribute continuum, and Table 7 shows this:

Table.7. is a summary of the results of the Reliability and separation coefficients and their standard errors for persons and items according to the Andrich-Rasch estimation scale model

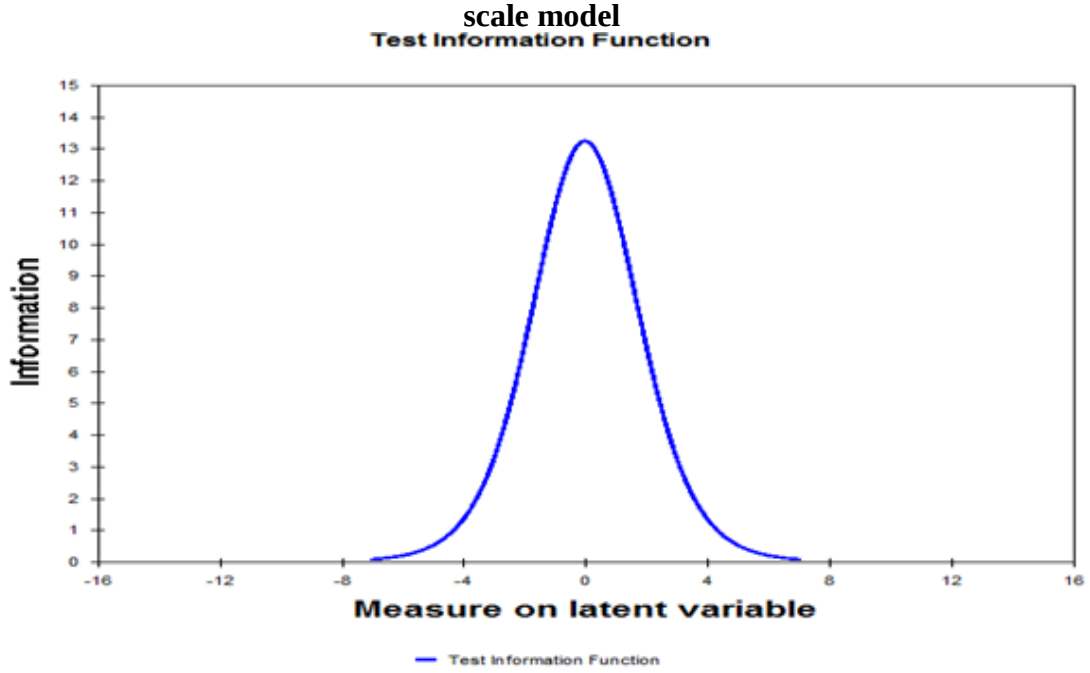
Reliability coefficient		Separation coefficient		Standard deviation		Root mean of standard Error	
Persons	**items**	**persons**	**Items**	**persons**	**Items**	**Real RMSE for persons**	**Real RMSE for items**
0.89	0.97	2.85	5.68	1.07	0.55	0.37	0.10
						persons RMSE model	**Items RMSE model**
0.90	0.97	3.03	5.78	0.07	0.55	0.35	0.10
Cronbach's alpha (or KR-20) Reliability coefficient, i.e. scale Reliability, is 0.91. SEM=3.25							

Source: winsteps v4.0.1 output.

It is clear from Table 7 that the value of the scale coefficient of separation is equal to (5.78), which is a value greater than (2), thus confirming the hierarchy of the scale items difficulty in the final calibrate, and that the lowest value for separating items is (3), and the lowest value of items Reliability (0.90), in order for these items to be sufficient to identify the continuum of the trait that it measures and the sufficiency of the sample of persons to confirm the hierarchy of items difficulty, and it is equivalent to the structural validity of the tool (Linacre, 2012). Also, the separation coefficient indicates the number of layers that make up the items difficulty of the scale, so the items spreads well on a logistic scale with high Reliability. This is what was recorded in the current study with the value (0.97) of the items Reliability coefficient. The value of the Persons Separation Index (GP) was also (3.03), which is a value greater than (2), as the lowest value for the separation of persons is (2), and the Reliability of persons (0.89) in order for the sample of persons to be sensitive to the distinction between high and low performance, In the current study, the values of both the separation and Reliability coefficients are considered acceptable, and through these coefficients (GI) and (GP) the Reliability coefficient was calculated for each of the items and persons, according to the following mathematical formula: ($R=G^2 /1+G^2$) (Andrich, 1982, pp.95-104), where (G) denotes the coefficient of separation, and (R) denotes the coefficient of Reliability. The values of class indicators range from (0) to infinity, and higher values are better. The value of the items Reliability coefficient was (0.97), which is a high value and indicates the adequacy of the sample of persons in separating the item, and thus in defining the continuum of the attribute measured by these items, which is independent of the length of the scale. The value of the reliability coefficient for persons was (0.90), with a standard error of (3.25) for the scale as a whole, which is acceptable, which indicates that the range of the trait is wide, that is, there is a good variance for the traits of persons, which is independent of the sample size. It should be noted that the value of the Reliability coefficient of persons in the "Rasch" model is approximately equivalent to the Reliability of the test in the classical theory Alpha Cronbach (α) or Couder-Richerson (KR-20) (Linacre, 2012), which represents the minimum reliability coefficient (Julian, 1988). .

Another indicator that indicates the Reliability of the scale using item response theory is the Test Information function, which helps determine the amount of information that a single item provides about an persons, by determining the shortest height of the curve representing the information function for that single item on the ability continuum, and thus it can be determined which test item best measures the variable to be measured (the trait) at specific levels of ability. The function of the scale information is an amount that is inversely proportional to the standard error of the measurement. Fig 6 shows the value of the information provided by the Negative Emotional Scale.

Fig.6. The information function of the scale according to the Andrich-Rasch calibrate scale model



Source: winsteps v4.0.1 . output

It is clear from Figure 6 that the value of the information function of the overall scale was as large as possible at the logit (0) attribute level, meaning that the scale gives more effective information for persons with a medium trait, while the values of the information functions provided by the scale were as few as possible at the high and low trait levels. It was also found that the value of the information function increases gradually with the increase of the attribute and reaches its maximum value when the value of the attribute is zero logit, or close to zero, where it reached (0) logit, and the value of the standard error of measurement is (0.27), and the maximum amount of information is (13.24), when the raw score (30), that is, the amount of information increases with a decrease in the standard error, and thus the value of the scale Reliability coefficient increases with it.

7- Conclusion:

Through the results of our current study, we have eliminated 12 items from the scale because they do not fit the scale model of Andrich estimation, which is derived from the Rasch model. The study also concluded, through the results of the validity of the factorial analysis of the main components of the residuals using the Rasch model, that the scale has one dominant dimension, which is measuring the trait of negative emotional reactions (depression, anxiety, psychological stress) among secondary school students, and a good structural validity according to the "Wright" map. We also recorded high values for the indicators of separation and Reliability of persons and items. While the process of evaluating the four alternatives of the scale resulted in the integration of the first and second alternatives together, and thus the scale became composed of three response alternatives, and we recorded the maximum value of the amount of information at zero logit for difficulty.

Therefore, through the findings of this study, we must draw the attention of specialists in the field of diagnosing negative emotions to the importance of the singular response theory in calibrate scales for diagnosing mental disorders and Depression; Anxiety; Psychological pressure, urging them not to rely on raw scores only in statistical analyzes in diagnosing these disorders, but rather to rely on the results of models of the item response theory to obtain more accurate results, and thus access to objective and correct decisions based on correct scientific foundations. Thus, the current study has provided the Arab environment with an objective scale characterized by accuracy in measurement, thus making it easier for researchers and specialists to use it in the processes of diagnosing these disorders without the need to ascertain its psychometric properties, as the calibrate of persons is free from the impact of item, and the calibrate of items is free from the impact of persons, which makes us

overcome the problems related to legalization, which gives its results a kind of credibility and the possibility of generalizing its results to persons with similar characteristics.

References:

Beeker, Frank (2010). *Foundations of the Individual Response Theory* (Translation by: Al-Tariri, Abu Hashem, Sawsan). Saudi Arabia: Scientific Publishing and Printing Press, King Saud University.

De Ayala, R.J. (2017). *Theory and practice in paragraph response theory*. (Abdullah Zaid Al-Kilani and Ismail Salama Al-Bursan, translator). Riyadh: King Saud University Press.

David J.A. Dozis and Keith S. Websen (2015). *Prevention of anxiety and depression: theory, research, and practice*. (Translation and submission: Abdullah Al-Sayed Askar). Cairo: The National Center for Translation.

Ali Asghari et al. (2008). *psychometric properties of The Depression Anxiety Stress Scales (DASS21) in non- Clinical iranian Sample*.International journal of psychology. vol2,no2, 82-102 ; winter 2008

Andrich, David. (1982). *An Index of Person Separation in Latent Trait Theory, the Traditional KR-20 Index, and the Guttman Scale Response Pattern*. Education Research and Perspectives, 9:1, 95-104.

Dominique, S. E. R. V. A. N. T. (2012). *Gestion du stress et de l'anxiété*, 3e éd. Copyright © 2012 Elsevier SAS.

De Ayala, (2009). *The Theory and Practice of Item Response Theory*. The Guilford Press A Division of Guilford Publications, Inc. 72 Spring Street, New York, NY 10012. www.guilford.com

Linacre J.M. (1999). *Investigating rating scale category utility*. Journal of Outcome Measurement,3:2, 103-122.

Linacre,J.M.(2006). *Auser's guide to WINSTPES MINISTEP Rasch model computer programme manual 3.92.1*.Retrieved from http://www.Winsteps.com.

Linacre,John ,M.(2012).*A User's Guide to Winsteps Ministep Rasch- Model Computer Programs*. Winsteps.com

Wright, B. & Masters, G. (1982). *Rating scale analysis*. Chicago, IL : MESA Press.

William j; Boone et al. (2014). *Rasch Analysis in the Human Sciences*. Spinger Dordrecht Heidelberg New **York London. http: extras.springer.com**

How to cite this article by the APA style:

Sonia AIOUADJ, Rachid ZIAD (2023). Use of Rasch-Andrich's Rating scale Model in the Calibrating of the Depression, Anxiety and Stress Scale (DASS42) on high school students .*Journal of Psychological and Educational Sciences*. 9 (2). Algeria: El-Oued University.. 133-147