The Use of Psychometric Scale Theory in Formulating Gilliam Scale GARS-3 for Diagnosing Autism Spectrum Disorder

Inst. Abbas Fadhil Abdulwahid
University of Baghdad/ College of Arts/Iraq
abbasfadhil@coart.uobaghdad.edu.iq
Prof. Dr. Muhammad Anwar Mahmoud
University of Baghdad /College of Education Ibn Rushd /Iraq
Mohammed.a@ircoedu.uobaghdad.edu.iq
Prof. Dr. Khalid Jamal Jasim
Khalid.jamal@ircoedu.uobaghdad.edu.iq
University of Baghdad /College of Education Ibn Rushd /Iraq

Received: 2/6/2022 **Accepted: 5/7/2022** **Published: 15/12/2022**
Doi: https://doi.org/10.36473/ujhss.v61i4.1998

Abstract

The present paper aims to formulate Gilliam Scale GARS-3 for diagnosing autism spectrum disorder (ADS) according to the use of psychometric scale theory. In order to achieve the aim of the paper, the researchers have followed the scientific steps in formulating the scale and verifying the credibility of the translation via checking the accuracy and subjectivity in transferring all the thoughts of the scale. Then, an expert in Arabic language has checked the scale which consists of (58) items of six dimensions. The scale was applied to (500) students inside the governmental and the private centers and on a sample that was chosen according to the random class style includes the parents and trainers in Baghdad government. Likewise, to verify the psychometric features of the scale, three controls have been acknowledged for selecting the items. The controls are: the recognizing factor control through the use of the two extremist group style of the total score. Item credibility criterion that depends on an indicative associative relation between the score of the item and the total scores of the scale by using Pearson's associative factor. The ratio of the items saturation depending on the factor explorative analysis. The results of the analysis show that validity of all the items. The credibility of the scale has been verified by two ways, namely:The explicit credibility: through showing the instructions of the test and its items to a group of experts at the department of educational and psychological science. The construction credibility: through the explorative factor analysis: finding out the stability of the scale according to Alfa Cronbach, and the stability factor scored (0.941). In the light of the procedures followed by present paper, the researchers have reached to some recommendations and suggestions.

Keywords: autism spectrum disorder, Gilliam scale GARS-3, traditional scale theory.

Statement of the Problem

Designing scales, measures and tests together with their development should be based on organized and rational approach (method).To achieve this process, there should be a framework supporting this method as well as authentic and bona fida techniques to reach high levels of reliability and validity in measurements and new testing (Rattray & Jones,2007, p.234) due to the difficulty of delimiting many psychological concepts and constructs. For this reason, one cannot rely on one measure tool (instrument).So , it is advisable to adopt different measure strategies by using suitable and divergent measures so that we can achieve better and higher profile for the construct being measured.(Alaam,2000,p.478).Thus , the accurate diagnosis of ASD is regarded as one of the most important issue according to specialists and trainers through conducting appropriate therapy programme via its application together with its assessment of results properly carefully where there are overlapping symptoms with other disorders. However, these factors led to disagreement among specialists in diagnosing ASD. Accordingly ,these differences had positive consequences in arriving at objective diagnosis based on scientific true grounds (Shreiban , 2010,p.60).As a result, choosing Gilliam Scale (GARS-3) for diagnosing autism spectrum disorders in accordance with psychometric scale therapy pursuant to developments that occur on the measurement of GAR-3.Since GAR-3 has been subjected to many changes and alterations if compared with the first edition as well as sixteen items have been mentioned only out of the first editions with adding forty two new items. Moreover, this measurement has been delimited to reflect the changes that occurred on ASD. It is found that DSM-5 (model had been excellent and superior over DSM-4 (model).

The sample was suitable and stable for the age and gender even with these individuals who has advanced symptoms of ASD. In addition to these changes, the measurement has been subjected to analysis of correlation coefficients and theoretical and scientific reliability for its subsidiary standards. The measurement has been applied on collected samples in (2010-2011) which have been compatible with existed demographic features in the United States of America, the year aforesaid.

This measurement is easily applicable measurements; its results can be reliable and verified. For this reason, it is regarded as one of the modern measures which reflects in its construction the definition of ASD according to both DSM-5 AND autism community.

One of the strength points of the measurement is that it attempts to link assessment by required intervention by means of implication of recommended guiding objectives through its implementation of each item of measure (Gilliam,2013, p.14)

Accordingly, the problem of the current research paper has been followed –up providing and conducting diagnostic tool (instrument) of autism spectrum disorder especially the Iraqi environment lacks a diagnostic standardized tool to diagnose individuals carefully in specialized centers and institutes of re-habilitation and care looking after those individuals who are affected with ASD .Therefore, there should be dire need for such a diagnostic tool that is distinguished by psychometric characters that can help diagnose individuals carefully with less effort and cost.

Significance of the Study

Researches on ASD have been achieved a great advance in the last two decades in terms of figuring out the rates of diffusion together with laying down measures of careful assessment in accordance with modern diagnosis standards. Due to the remarkable rise of rates of diffusion of ASD, some inquiries have been raised concerning the accuracy of employed methods together with the recrudescent focus on advance diagnosis techniques. Besides , the wrong diagnosis of ASD is deemed as real threats regardless of achieved progress in this field.

On the account of the diagnosis of ASD and its delimitation are equally problematic. So, adopting psychometric approach in assessment and employing accurate standards are momentous in parallel with developing measures of assessment of ASD (i.e. development of actual diagnosis standards).That is why diagnosis standards of ASD have been laid down more clearly even though there is much constant debate around it (APA,2013,PP.3-6).

Generally, diagnosis of ASD has not completed until after two or three years have passed out of the child's age. This was due to the misgivings concerning marking and incorrect diagnosis of children. Thus, identification of children who suffer ASD as well as the early intervention intensively in (3-7) pre-school years lead to improvement of the results of most children who suffer from ASD since the early diagnosis and intervention can facilitate the educational planning .This process can be done through every day routine observation, ASD –related examination, identification of ASD type carefully where profound and meticulous assessment and diagnosis are quite important in delimiting interventive disorders and accordingly defining the suitable therapeutic program for ASD (Fillpek et al , 2010,p.3)

There is a general agreement on the fact that the early diagnosis of ASD (including pre-childhood in extreme cases) is very important issue to achieve optional results and laying down suitable early therapeutic program for children due to the fact that the early diagnosis and examination result in an early intervention for those children which can help reduce family stress. Also, it can increase skills of confrontation and developing social adaptation for children in the future (Wing et al, 2010, p.5)

One of the main concerns is to improve the standards of social sciences. For this reason, psychometricians have attempted to take great care of excellent syllogism, quantitative techniques and employing scientific experimental methods (Alaam,2000, p.16).Tests and measures are tools that have been designed to be used to take human decisions .Nevertheless, in our complex and multidimensional society , thousands of decisions have been made each single day and these decisions which involve certain assessment of psychometric characteristics for individuals.

Some of the decisions made by some people have been taken to regulate their private lives and some of these decisions are diagnostic and some of them are evaluative (AdulRahman, 1998, p.325).The process of reconstructing tests and measures is considered one of the basic technical process which students of measures should a train and know in psychology (Alaam,2000, p.479). It is

important to conduct reliable and authentic techniques in setting up approved psychometric characteristics and any default in constructing adequate and efficient measure leads to a difficulty in interpreting results and findings and this will affect the educational practices (Rattray & Jones,2007,p.234)

In the light of what has been said earlier, the significance of the current study contributes a great deal to enrich the libraries of psychometric tests by highly authentic and accurate measure which is characterized by validity, objectivity, and reliability. Also, it can represent all psychometric characteristics which resulted in achieving objective results and accurate and precise diagnosis.

Purpose of the Study

The current research paper aims to use psychometric measure theory in formulating the Gilliam Autism Rating Scale (GARS-3) to diagnose Autism spectrum disorder.

Limits of the Study

The following research paper is limited to study the following:

1. The original version of Gilliam Autism Rating Scale (GARS-3) to diagnose autism spectrum disorder.
2. Children of centers and institutes (private and public) in Baghdad governorate for both genders.

Definition of the Terms

Psychometric Theory:

Alaam (2001) defined psychometric theory as "balancing individual performance in relation to Peers performance and explaining its score only in the light of this group reference since the score which the individual obtains is meaningless and unexplainable if it is not subjected to this standard due to the fact that we can not balance among individuals on the raw scores which they obtain out of the test and individuals have been classified according to their relative positions in terms of capabilities and different features. (Alaam, 2001, p.16)

Gilliam Autism Rating Scale -3 to diagnose Autism spectrum disorder:

"It is a tool designed to assess and evaluate individuals who suffer from Autism spectrum disorder and other acute behavioural disorders.It present information that can help diagnose disorders ".(Gilliam,2014,p.4)

Autism spectrum disorder = (ASD)

Diagnostic and statistical manual of mental disorders(-5) defined ASD as :

"Inability to interaction and social communication with decreased sharing of interests with others , difficulty appreciating their own emotion ,lack of proficiency with use of non-verbal gestures, a version to maintaining eye contact as well as difficulty in adaptation and making friends and taking parts in playing (DSM-5,2013,P.50)

Theoretical Framework and Review of Previous Studies

By viewing some reviews of literature of the psychometric measurement, it is found that there are two general theoretical trends under which psychometric theories fall that attempt to explain in finding scientific methods and techniques to

reach high level of accuracy , precision and objectivity in measurement .The former is called the classical trend represented by classical psychometric test theory. The latter is called contemporary trend or contemporary theory, due to the fact that formulating the scale is based on the psychometric theory; researchers will present the theory briefly.

Classical Psychometric Test Theory

Classical Psychometric test theory is considered one of the oldest theories of psychometric measure and it is more commonly used since (1904) up to now. Charles Spearman was one of the founders of this classical test theory .It is often called a true score theory and error free score. Classical test theory assumes that every person has a true score on an item or a scale if we can only measure it directly with no error. Spearman confirmed that the individual's score resulted from his/her score plus error score.(Krus & Fuller, 1982:p.37)This theory interprets the variation which occurs between the frequencies of measurement for one individual for being an error score which relates to irregular factors that overlap with true score which reflects what the individual has out of measured variable and other variables (Chiselli et al 1981,p.195)

This theory is based on a basic assumption at the time of formulating educational and psychological tests with analyzing its items. This assumption states that "distribution of individuals scores in the measured characteristic via test is through moderate distribution which is influenced by individual sample and characteristics of test item sample since individual differences are based on peer group. This means that the achieved results out of characteristic measure performs to group standard which they belong to .This type of measure is called (Norm-referenced) (Culler, 1966, p. 272). This theory is based on basic (key) postulates (axioms) in psychometric measure with some hypotheses out of it.

Individual performance can be measured and evaluated since individuals have similar and different behaviours with each other and these behaviuors can be measured and transformed from descriptive form and quantitative one by using numbers according to specific rules and thus we can predict subsequent stages of these behaviuors.

1. The individual performance is an indication that his/her characteristics since each behaviuor or performance which resulted from one characteristic or group of characteristics individuals have or possess .Accordingly, one characteristic or feature gives more than one performance and one performance resulted in more than one feature. Thus, we can see that there is a relationship between the characteristic and performance in terms of quantity and quality which influenced the nature of used tool in terms of construction , meaning and interpretation.
2. The relationship between the characteristic and performance differs from one individual to another due to individual differences among person. This is

   The key principle which psychometric measure process is based on.(Abdulrahman,1983,p.80-88)

3. The individual's score on psychometric measure consists of two scores. They are true scores and error score.(Weiner & Stewart,1984,p.57)

(A) Systematic Errors

Is the error which is resulted from systematic factors (i.e.,it is repeated systematically and it was the same impact in scores all individuals on the used measure and its quantity can be predicted so it does not constitute any problem of measurement and it does not influence reliability of measure .

(B) **Random Error** is a type of error which can't be predicted because it is random error and it is governed by coincidence and its consequences can be decreased by increasing the size of the sample since random errors cancel each other.(Magnusoon,1907,p.61-62)

The true score and error score which this theory stresses on have three hypotheses. They are:

1. The first hypothesis:

It is related to a true score (i.e. the individual has stable characteristics which remain stables during time and this hypothesis should be taken for granted and if the true score of the individual changed continually, there is no stability in his/her characteristics as for the noticed score changed continually since the range of error changed from application to another.(Ghiselli,1964,p.221)

2. The second hypothesis

This hypothesis is related to the error score (i.e. errors are randomized complete error meaning that they are not systematic since these errors distributed moderately and their average will be zero and equal variance for all individuals and the relationship between true score and error score is that the true score is stable and error score is changeable.(Malham,2000,p.254)

3. The third hypothesis shows that the addition (totally) of true score and error score resulted in noticed (apparent) score which is changeable due to instability of error score.(Ghiselli,1964,p.221)

According to this hypothesis, the total variance which represents the variance in the notices score is divided into two parts: the true variance which reflects true differences among individuals and error variance which represents random error in measure. As for validity in accordance with this theory, it represents the ratio of true variance related to the measured characteristic and to the noticed variance (Auda,2000, p.339-340)

The true variance which measures the validity is included either in the characteristic or the function intended to be measured or according to the nature of the sample or population which intends to study certain characteristic (Al-Imam et al,1990, p.124) Given that the true test measures true differences among

individuals in the characteristic. So, it is advisable to benefit from laying down decisions that related to limited aims and objectives.(Alam,1989,p.209)

Models of Classical Psychometric Test Theory:

Classical psychometric test theory includes two basic models. The first is called Domain sampling model. The second is called parallel tests model.

1. **Domain Sampling Model**

   This type of model is considered one of most commonly used and classical models in formulating tests and educational and psychological measures. It is based on the assumption that there is a comprehensive domain out of items universe can be selected random samples out of these items which the test or measure should be included. This model aims to access and evaluate the score of items universe.

2. **Parallel Tests Model**

The parallel tests model assumes the actual parallelism for the two tests according to the random observations since the two test are parallel if the mean and standard deviation equalized with their scores and they are equally related with a group of true scores (Alam, 2004,p.49-50)

Diagnostic of Autism Disorders (ASD)

The early examination and diagnosis leads to early intervention for children who suffer from disorder which help delimit the family stress, increase skills of confrontation, develop social adaptation in the future. Thus, institutions and institutes that are concerned with (ASP) that the early diagnosis decreases the parents' doubts and helps them delimit options of suitable learning, as well as facilitate in educational guidance whereas the belated therapy is seen as a default in the early diagnosis. So, the early care is defined as a group of interventions that target children from childbirth to childhood (6 years old) and family as well as environment. It also aims to respond to temporary and prominent needs of children. Accordingly, a group of specialists have been established such as psycholingists, educational-psychologists, speech therapists, physical therapists, psychotherapists, and neurologists.(Bakers et al, 2014, p.2)

In most cases, ASD has not been diagnosed after two or three years.Passing out of children's ages. This stems from the fears concerning marking or diagnosing children wrongfully. For this reason, diagnosing children who suffer from ASD and starting early and intensive intervention during pre-school years (3-7)leads to improvement of the results of most children .Accordingly, the early diagnosis and intervention facilitate educational planning through observation, and examination.(Filipek et al,2010,p.5)

In order to diagnose Autism spectrum since childhood, there should be a group of standards for such diagnosis:

1. Qualitative disability in terms of social interaction , abnormal behavioural characteristics, loss of central over eye contact and facial gestures which lead to weakness in personal relationship with environment and surroundings.

2. Qualitative disability (defect) in verbal communication together with weakness in speech and language acquisition.
3. A limited group activities and interests which include repetition movements and recurrent behavioural patterns such as (Fluttering , Clapping , codependency , and getting used to daily routine (Al Zuraifat , 2004,p.127)

Literature of Previous Studies

- **Abdulrahman & Khalifa study (2004)**

In this study, Abdulrahman & Khalifa Arabized Gilliam scale for diagnosing autism spectrum for those individuals whose ages range from (3) to (22) years old and those individuals suffer from autism spectrum disorder and those who suffer from a cute behaviuoral problems. The purpose of this scale is to help specialists to diagnose autism and the items of the scale have clear and strong face validity since they are based on defining autism which is conducted by the American Association of Autism .Besides, they are characterized by diagnosing standards which are presented by the American Psychiatric Association in the diagnostic and statistical mental of mental disorders (fourth edition).This scale consists of four sub – dimensional that describe limited behavior and it can measure each dimensions by 14 items and the four dimensions are patterned behavior of communication & interaction. The raw scores can be achieved and transformed to standard scores and percentages as well as the total of standard scenes into Autism quotient .Accordingly, Gilliam autism scale has excellent psychometric characteristics. (Abdulrahman & khalifa , 2004,p. 270)

- **Theeb 2013 Study**

The psychometric characteristics of Syrian image of Gilliam scale (GARS-2) to diagnose autism spectrum disorder. The study aims to identify the psychometric characteristics (second edition) of Syrian image out of Gilliam scale to diagnose autism spectrum disorder (GARS-1) .To achieve the objectives of the current study, the researcher adopted the descriptive –analytical approach and translated the scale from English into Arabic. The scale is applied on a sample consisting of (240) individuals who have been selected intentionally distributed into (176) individuals who suffer from autism spectrum and (31) individuals who suffer from mental disabilities and (33) who are normal individuals.

Results show that these scales has content validity and discriminative validity between a group of autism disorder individuals and mental disabled individuals and normal children (Theeb ,2013, p.101-135)

- **Samadi & Mccnokey (2014)**

The study aims to assess the use of autism spectrum disorder diagnosis scale (GARS) which is commonly used in the Western society and Iran. This scale has been translated into Persian language and applied on a sample of guardians of (658) children. After diagnosis , it is found that (442) persons who have been diagnosed as having autism spectrum disorder and (112) individuals who have mental

disorder and (102) normal children. Accordingly, the subsidiary scales have been evaluated and compared with the three groups. Results show that the three branches have high internal consistency and those individuals who suffer from autism have been distinguished from the other two groups and the defining cut score increases out of sensitivity of the scale(Samadi & Mccnokey, 2014,p.2-4)**.**

Firstly, Research Methodology

Researchers conducted descriptive approach in the current study due to the fact that this approach is most commonly used and it is indispensible and consequently, any phenomenon requires a detailed description and value (Dawood & Abdulrahman, 1990, p. 159).

Secondly, Research Population

The population of the current study consists of a group of students of private and public centers from both genders (males and females) subsidiary to Council of Baghdad Governorate and Ministry of Labor (Al-Karkh/Al Rusafa) for the academic year (2017-2018).So, the population size amounted to (4422) diagnosed male and female students distributed on the public and private centers (Al Salihy,2020, p.10)

Thirdly, Basic Research Sampling

It is worth mentioning that there is no limited number out of the original size of population that can be applied on all studies (Doweidry,2000, p.307)
In the current study, two samples have been selected as follows:

1. Clarity sample of items Instructions
   The main purpose is to identify the range of clarity instructions and items among sample individuals. The sample amounted to (60) out of trainers and guardians.
2. Sample of Statistical Analysis:
   It amounted to (500) out of trainers and guardians which has been applied on a sample of males and females.

Fourthly, Research Tools

The research tool is a used method of collecting data to achieve the objectives of the current research paper. Thus, in the present research paper, the researcher has adopted an adequate research tool according to the standards of modern diagnosis(DSM-5).Accordingly, the researchers have obtain the original version of Gilliam test (GARS-3)which was published in 2013 in the United States of America.

Scale Description

It is a reference tool designed to diagnose autism spectrum disorder (ASD) for individuals whose ages range from (3-22) years old.GARS-3 consist of three different components or parameters. The tester manual includes (50) application forms with pamphlet of educational objectives for those who are afflicted with

autism. The tester manual contains administrative procedures of assessment – interpretations of those producers and measures as well as technical information of the test. The tester fills in the application form which involves demographic information of the medical case and item answer of the scale (GARS-3).
Besides, it contains closely-related information of autism. The scale also contains booklet of educational objectives which can help those afflicted with autism together with specialists to develop educational objectives and behavioral interventions in accordance with the outputs of this scale.
In harmony with the definition of autism population (2012) and autism disorder according to (DSM-5) and APA (2013).
The purpose is to help specialists who can diagnose autism carefully.
Moreover, it can help lay down suitable therapeutic programs .This scale has excellent psychometric characteristics which enable those specialists to complete this scale during a period ranging from (10 to 15) minutes (Gilliam, 2014)

Fifthly, Procedures of Formulating the Scale
Validity of Translation
Translating of the scale is regarded as one of the most important issues due to the fact that the validity of the scale score is based on the accuracy of translation together with adaptation of the test to the new environment (Giert, 2000, p.281)
Therefore, the scale has been rendered from English into Arabic to verify the authenticity of its Arabic language as well as the validity of its instructions together with its items for logical and statistical applications.
Accordingly, the researchers followed a group of procedures to achieve the intended purpose as follows:
Translating the test from English into Arabic and then re-rendering the test from Arabic into English, so as to find equivalence between the original scale and the translated one .Thus, we can measure its validity and accuracy .Given that the scale has been submitted with its Arabic version to an expert in Arabic language to verify the validity and integrity of translating process.

The Logical Analysis of Scale Items
The scale should apparently measures what has been intended to be measured. It means that during examining the test, the specialists draw conclusion that the scale has measured correctly (Al-Batch and Abu Zina, 2007, p.128). To verify the matching of items for the charcateristics. The scale has been submitted to a group of specialists in educational and psychological sciences to verify the appropriateness of the scale .The experts have approved 100% the validity of the scale items.

Clarity of Instructions and Understanding the items
The researcher, for identifying the range of clarity of instructions and understanding the items together with the time of response, has chosen
Random sample of research population that comes to (60) individuals distributed equally between males and females. It appeared that instructions of the scale and its

items are clear. The elapsed time of the response ranges from (10-15) minutes. Given that the time of reading of general instructions has not accounted within a given time.

The Statistical Analysis of Items

To verify the psychometric characteristics of the scale which include discriminant force and items validity. Students have been selected randomly in each centers, institute as well as school of autism and difficulties of learning and mental disability (by means of random selection).The number of students has amounted to (500) males and females and this number is suitable according to (Nunnally, 1978). Nunnally suggests that the size of the sample of items analysis should be between (5-10) individuals for each item to reduce the chance effect (Nunnally, 1978,p.262).

For the purpose of carrying out the statistical analysis of the items . Scores have been arranged in descending order and then the two extremist group have been specified with 27% out of individuals in each group. Thus, the number of higher group is (135) individuals and (135) in the lower group. After using T-test for two independent samples for testing significance of statistical differences among the two means of scores of the two groups for each item of the scale. The T-value calculated refers to the discriminant force of the item (Ellis ,1976,p.50).

It is found that all items have discriminant force as illustrated in table (1)

Discriminant Force of the Item of the Scale

Group	N.	Mean	Std. Deviation	t
VAR00001 Higher	135	2.0963	1.03574	8.522
Lower	135	.9481	1.17383	8.522
VAR00002 Higher	135	2.2222	1.09018	10.446
Lower				10.446
	135	.9259	.89948	14.078
			.94999	14.078
VAR00003 Higher	135	2.2222	.98345	12.274
		.9259		
Lower	135		.94999	12.274
			.95632	11.991
VAR00004 Higher	135	2.1630	1.08025	11.991
Lower	135	.5778	.78593	13.521
VAR00005 Higher	135	2.1333.	1.00149	13.521
		.6889		
Lower	135			
		.8148	.67858	10.468
VAR00006 Higher	135	2.4148	1.19368	10.468
Lower	135	.9333	.79994	10.761

Group		N.	Mean	Std. Deviation	t
VAR00007	Higher	135	2.4815	1.30136	10.761
	Lower	135	1.2444	.96351	14.295
VAR00008	Higher	135	2.5037	.99457	14.295
	Lower	135	1.0889	1.02349	16.087
VAR00009	Higher	135		.78417	16.087
	Lower	135		.93999	13.507
VAR00010	Higher	135		1.05708	13.507
	Lower	135		.87234	17.030
VAR00011	Higher	135	2.4000	.92098	17.030
	Lower	135	.6963	1.13821	13.234
VAR00012	Higher	135	2.1852	.88829	13.234
	Lower	135	.4000	1.03781	12.569
VAR00013	Higher	135	2.2667	.94527	12.569
	Lower	135	.6222	1.02214	9.747
VAR00014	Higher	135		1.05052	9.747
	Lower	135	.5259	.97554	13.265
VAR00015	Higher	135	2.1333	.90408	13.265
	Lower	135	.4889	1.01061	12.802
VAR00016	Higher	135	.6222	1.00496	12.802
	Lower	135	2.0000	1.06130	6.461
VAR00017	Higher	135	.7704	1.15948	6.461
	Lower	135	2.0593	1.09337	12.177
VAR00018	Higher	135	.5407	.75508	12.177
	Lower	135	2.1259	1.08729	11.406
VAR00019	Higher	135	.5556	.78860	11.406
	Lower	135	1.5778	1.04768	9.931
VAR00020	Higher	135	.7037	1.04815	9.931
	Lower	135	1.7926	.95782	11.097
VAR00021	Higher	135	.4000	1.05791	11.097
	Lower	135	1.7630	1.04116	12.819
VAR00022	Higher	135	.4444	.90022	12.819
	Lower	135	.5630	.96872	11.471
VAR00023	Higher	135	1.9778	.90847	11.471
Lower		135	.6148	.76497	11.293
VAR00024	Higher	135	1.9259	1.13870	11.293
Lower		135	.4074	.90163	7.096
VAR00025	Higher	135	1.9037	1.12698	7.096
	Lower	135	.5926	.90493	14.166
VAR00026	Higher	135	.9037	.89942	14.166
	Lower	135	2.0889	1.07116	11.574
VAR00027	Higher	135	1.2074	.92134	11.574
	Lower	135	2.1556	.99240	12.637

Group		N.	Mean	Std. Deviation	t
VAR00028	Higher	135	.6000	.91302	12.637
	Lower	135	1.9037	.80915	11.543
VAR00029	Higher	135	.4963	1.09862	11.543
	Lower	135	1.9852	1.11336	14.580
VAR00030	Higher	135	.5185	.79745	14.580
	Lower	135	2.1778	1.02576	11.993
VAR00031	Higher	135	.8222	.84295	11.993
	Lower	135	.3630	1.06868	9.637
VAR00032	Higher	135	1.8074	.98849	9.637
	Lower	135	.4370	.80545	14.833
VAR00033	Higher	135	1.8519	.93734	14.833
	Lower	135	.6444	.92493	10.205
VAR00034	Higher	135		1.03973	10.205
	Lower	135		.98395	8.473
VAR00035	Higher	135		1.08219	8.473
	Lower	135	2.0889	1.04021	12.056
VAR00036	Higher	135	.5111	.93575	12.056
	Lower	135	1.9481	1.13549	7.743
VAR00037	Higher	135	.7259	1.15948	7.743
	Lower	135	2.0444	1.08907	10.312
VAR00038	Higher	135	.9778	.98569	10.312
	Lower	135	2.0074	.98704	7.409
VAR00039	Higher	135	.5556	1.11212	7.409
	Lower	135	1.7852	1.10943	13.439
VAR00040	Higher	135	.7037	.79363	13.439
	Lower	135	.6074	.98704	10.374
VAR00041	Higher	135	1.1556	1.14257	10.374
	Lower	135	.4000	.86948	12.186
VAR00042	Higher	135	2.1037	1.06805	12.186
	Lower	135	.7556	.98148	10.327
VAR00043	Higher	135	2.3185	1.24269	10.327
	Lower	135	.8741	.75968	4.656
VAR00044	Higher	135	2.4296	.92693	4.660
	Lower	136	1.0222	.75910	10.889
VAR00045	Higher	135	1.2222	.73668	10.887
	Lower	136	.7426	.97537	10.766
VAR00046	Higher	135	.5735	1.12544	10.771
	Lower	136	.9926	1.03893	11.942
VAR00047	Higher	135	2.3481	1.02219	11.941
	Lower	136	.8529	.90829	5.648
VAR00048	Higher	135	1.1037	.75954	5.644
	Lower	136	.5294	1.12103	11.172

Group		N.	Mean	Std. Deviation	t
VAR00049	Higher	135	2.2667	1.11389	11.171
	Lower	136	.7500	1.26482	10.202
VAR00050	Higher	135	2.1852	1.13085	10.198
	Lower	136	.6985	1.00348	8.666
VAR00051	Higher	135	1.5778		8.663
	Lower	136	.5662	.91653	7.601
VAR00052	Higher	135	2.1037	1.28292	7.601
	Lower	136	.9191	1.28268	8.979
VAR00053	Higher	135	2.4963	82746	8.989
	Lower	136	1.4338	1.10018	4.745
VAR00054	Higher	135	1.4148	.78593	4.748
	Lower	136	.9191	92740	8.736
VAR00055	Higher	135	1.7778	.75968	8.742
	Lower	136	.8824	.91941	7.502
VAR00056	Higher	135	2.0148	.47293	7.520
	Lower	136	1.2941	1.01191	8.493
VAR00057 VAR00058	Higher Lower	135 136	2.2963 1.2059	11.04435 .06887	8.494

Validity of Items

Coefficient of validity of items has been extracted experimentally in the psychological measures out of extracting correlation coefficient with external and internal criterion. It is the most important due to the logical validity and experimental validity refers to the range of correlation of construct content of the characteristic with each other (Abdulrahaman,1983, p.414-415).

To verify the validity of the scale items, the researchers conducted the total score of the internal criterion through which coefficient of scale items validity have been extracted out of the scale items. In case of unavailability of external criterion(Anatasi,1988,p.211).So, Pearson correlation coefficient has been used in each item score and the total score. After achieving results with weighing coefficients of calculated correlation in tabulated value of correlation coefficient showing that all items are significant statistically with level of (0.05) as shown in table (2):

Table (2) Coefficient the Validity of Scale Items –Autism Spectrum Disorder

SN.	Correlation	SN.	Correlation	SN.	Correlation	SN.	Correlation
1	0.430	16	0.540	31	0.549	46	0.508
2	0.482	17	0.511	32	0.452	47	0.485
3	0.555	18	0.351	33	0.436	48	0.524
4	0.519	19	0.480	34	0.522	49	0.312
5	0.537	20	0.462	35	0.474	50	0.488
6	0.533	21	0.446	36	0.432	51	0.467
7	0.523	22	0.454	37	0.518	52	0.422

SN.	Correlation	SN.	Correlation	SN.	Correlation	SN.	Correlation
8	5.557	23	0.493	38	0.393	53	0.393
9	0.541	24	0.495	39	0.439	54	0.277
10	0.565	25	0.529	40	0.390	55	0.277
11	0.562	26	0.360	41	0.520	56	0.412
12	0.610	27	0.526	42	0.475	57	0.453
13	0.570	28	0.489	43	0.506	58	0.459
14	0.538	29	0.490	44	0.497		
15	0.477	30	0.493	45	0.216		

Validity of the Scale

Validity is one of the most important features of psychometric characteristics that must be available in psychometric measures (Ebel,1972,p.435).because it is considered an indicator of the ability of the scale to measure (Harrison,1983,p.11). And it is based on face validity and construct validity. The face validity has been verified and approved. Specialists of educational psychological sciences on the validity and integrity of items .As for construct validity has been verified by means of:

1. Calculating discriminant force of scale items.
2. Relationship item scale by the total score of the scale with calculating correlation coefficient in each score and the grand total.

The Factorial Analysis

The factorial analysis is regarded as one of the important indicators in defining construct validity (Abu Hatab,1987, p.163) through which we can identify the range of internal consistency between items of the scale with each other and we can infer the factorial validity of the grand total of the scale if the correlation coefficient among items and the scale are very high (Samarah et al ,1989,p.113)

Correlation coefficient of the item can be reached by using factorial analysis which identifies the basic constructs of the phenomenon which are subject to measurement, so the factorial analysis is the powerful and strongest way to identify the validity (Aldulaimy,2004, p.98)

This type of validity helps us clarify the number of factors within a set of items as well as it helps us determine the nature of association among factors. Besides, it helps us determine which items are linked to which factor, which facilitates the interpretation of those factors (Al Nabhan ,2004,p.136)

The test has been subjected to factorial analysis by using SPSS to versify one-dimensional test and then the exploratory factorial analysis have been used by screen plot. Table (3) shows the factors, latent roots, interpreted variance and ascending cumulative frequency.

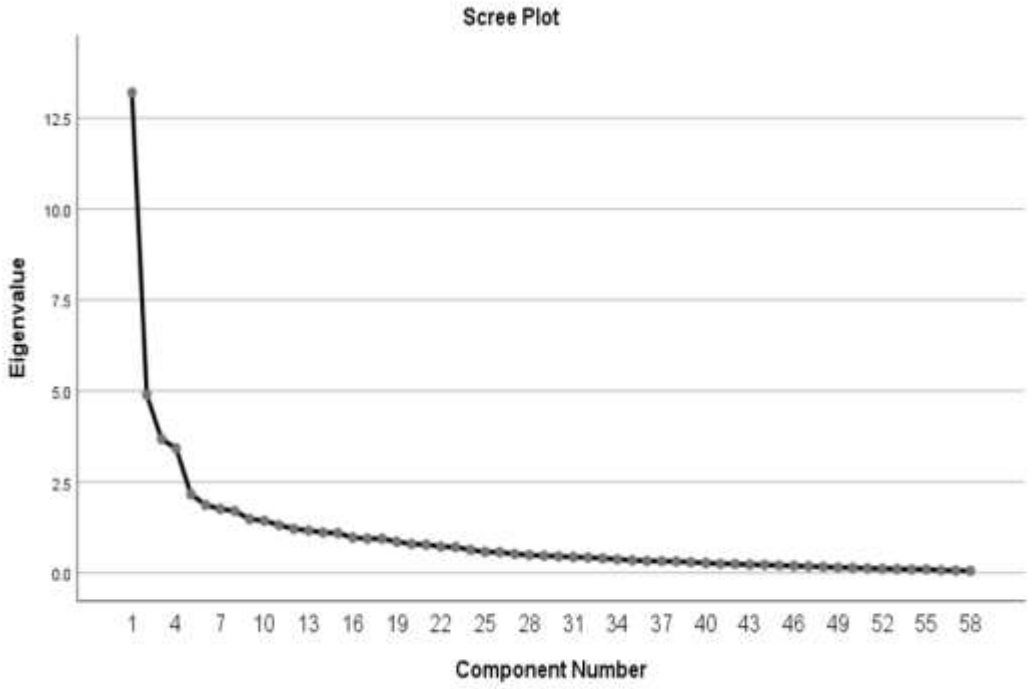
Table (3) The Latent Root , Ratio of interpreted variance, and ascending cumulative frequency:

Factor	Latent root	Interpreted Variance	Ascending Cumulative Frequency
1	13.202	22.762	22.762
2	4.897	8.443	31.205
3	3.679	6.343	37.548
4	3.418	5.894	43.442
5	2.160	3.723	47.165
6	1.865	3.215	50.380
7	1.756	3.028	53.408
8	1.698	2.928	56.336
9	1.475	2.543	58.879
10	1.433	2.470	61.349
11	1.311	2.260	63.609
12	1.207	2.081	65.690
13	1.161	2.001	67.692
14	1.107	1.909	69.601
15	1.092	1.882	71.483

Table (3) shows that there are fifteen factors with latent roots are above one and the interpreted variance of the first factor amounted to (23%) while the second factor variance comes to (8,44) where the ratio of the first factor to the second one is (2,69).Accordingly , the two conditions of one dimensional test have been achieved within (20%) out of the total variance and the ratio of the second factor to the first one comes to multiple fold.

As a result of what has been discussed earlier, all items are saturated on the first factor and this proved that the scale has a general factor. Figure (1) illustrates this fact:

Figure (1) Saturation of All items on the first factor



This figure shows that the distance between the first and second factor is greater than the distance between the second and third factor. Moreover, the distance among other factors is very convergent which indicates that there is one prevalent factor.

Scale Reliability

Reliability is one of the basic psychometric characteristic of psychological measures with regard to the fact that validity has priority over reliability since the valid scale is reliable whereas the reliable scale may not be valid. It cannot be said that each valid test is reliable (Al Imam et al ,1990,p.143).To calculate reliability , the scale has been applied on a sample consisting of (500) male and female students. Furthermore, reliability has been calculated by Cronbach's Alpha because this method is preferable to measure the reliability by measuring internal consistency between test items (Anastasi & Urbina,1997, p.95)

Accordingly , the reliability coefficient amounted to (0,941) table (4) illustrates this :

Reliability Statistics

Cronbach's Alpha	N of Items
.941	58

The Final Formula of the Scale

After verifying the psychometric characteristics of the scale represented by validity and reliability , thus the scale consists of six sub-fields distributed over a number

of total items of the scale which came up to (58) items distributed as follows: (13) items for repetitive and restrictive behaviors , (14) items for social interaction , (9) items for social communication, (8) items for emotional response, (7) items for cognitive style and (7) items for unadaptable speech, therefore , all these components are regarded as indicators of measuring autism spectrum disorder (ASD) and other acute disorders and all these items can be reliable and contributing in conducting important diagnostic decisions and they obtain correction (0,1,2,3).Diagnosis of disorder can be done on the original criterion of the test which can be adopted as important diagnostic criterion if they are distributed as follows:

- (54) degree and less (not carrying disorder)
- (55) – (70) degree (carrying disorder) at a simple level.
- (71) - (100) degree (carrying disorder at average level.
- (101) degree and more (carrying disorder) at intense level.

Statistical Methods

The researchers employed different statistical means in carrying out formulating autism spectrum disorder and in analyzing collected data out of research sample by using statistical package for social science (SPSS) and the statistical means are as follows:

1- Pearson correlation coefficient to calculate coefficients of items validity and to identify reliability by means of re-testing.
2- T-test for two independence samples to test the significance of difference between the higher group and lower group when calculating coefficient of discrimination.
3- Cronbach Alph equation has been used in calculating coefficient of reliability (internal consistency)
4- Factorial analysis by using varimax rotation.

Recommendations

1. Benefiting from the current scale as a standard of diagnosing children who are afflicted with (ASD) as well as medical diagnosis.
2. The possibility of using the current scale on normal schools students by other researchers since the scale has good psychometric characteristics.
3. Adopting the scale in classifying students of centers and institutes according to their whereabouts.

Suggestions for Further Research

1- Conducting a study aiming at setting up training programs to develop skills of communication for ASD children.
2- Conducting a study aiming at comparing between psychometric characteristics of the Iraqi image and the scale original image.

3- Conducting a study aiming at comparing between the classical theory and latent trait theory in psychometric characteristics of Gilliam scale (GARS-3) to diagnose ASD

References

- Abdulrahman M. and Khalifa , M. (2004) Manual of Parents and Specialists in Early and Intensive Behavioural Treatment of Autism Child.Cairo:Dar Al-Fakar Al-Araby.
- Abdulrahman, S. (1983) The Psychological Measurement (3rd edn.).Kuwait:Al-Falah Bookshop.
- Abdulrahman, S. (1998) Psychological Measurement .Kuwait: Al-Falah Bookshop.
- Abu Hatab, F. (1987) Mental Abilities .Cairo: Anglo-Egyptian Bookshop.
- Al- Dulaimy, H.R. (2004Setting up a Test of Shapes Inference Annals Journal of Intelligence Researches. Issued by unit of Intelligence Researches and Mental abilities. Issue (1) College of Basic Education. Al-Mustansirriya University.pp.69-104
- Al- Zureiqat, I. (2004) Autism Characteristics and Treatment.(1st edn.) Amman: Dar Wael.
- Alam , S. (1986) Contemporary Developments in Psychological and Educational Measurement .Kuwait: Kuwait University.
- Alam , S. (2000) Educational and Psychological Measurement and Assessment Fundamentals-Applications and Contemporary Orientations.College of Education Al Azhar University ?(1st edn.) Cairo: Dar Al Fakar Al-Araby.
- Alam , S. (2001) Diagnostic Criterion –referenced Tests in Educational , Psychological and Training Fields. Cairo: Dar Al Fakar Al Araby.
- Alam , S. (2005) Response Samples of One Dimensional and Multidimensional Test Item with its Applications in the Psychological and Educational Measurement (1st edn.) Cairo: Dar Al Fakar Al Araby.
- Al-Batsh, M., Abo Zina , F. (2007) Scientific Research Methodology. Research Design and Statistical Analysis. Amman:Dar Al- Sirrah for publication, Distribution and Printing.
- Al-Imam , M. (1990) Measurement and Assessment .Baghadad: Dar Al-Hikma.
- Al-Nabhan, M. (2004) Fundamentals of Measurement in Behaviuoral Sciences (1st edn.).The University of Jordan: Dar Al Shorouk for Publishing and Distribution.
- AlRawy , F. and Hamad, A. (1999) Autism .The Vague Disability .Al-Duha. Qatar.
- American Psychiatric Association. (2013) Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Arlington, VA, American Psychiatric Association, 2013.
- Anastasi , A & Urbina , S. (1997) . Psychological testing, 7th edn. , New York , prentice hall .

- Anastasia , A. (1988) Psychological Testing , New York , 6th Macmillan Publishing.
- Auda, A. (2000) Measurement and Assessment in Educational Process. Jordan-Dar Al Amal for Publication and Distribution.
- Bakare M, (2012): Munir K, Aguayo P, Gaddour N, Öner Ö, Mercante M, Autism Spectrum disorder in Rey JM (ed) , IACAPAP e- textbook of Child and Adolescent Mental health.
- Culler , E.A. (1966) Studies in psychometric theory , *Journal of Expert of Psychology* , Vol , 9 , No.2 .
- Dawood, A. and Abdulrahman ,A.(1990) Educational Methodology. Ministry of Higher Education and Scientific Research .University of Baghdad.
- Doweidry, R. (2000) Scientific Research. Theoretical Fundamentals and Scientific Practices.Damascus:Dar Al Fakar.
- DSM-5. (2013). Diagnostic and Statistical of Mental Disorders, Fifth Edition: American Psychiatric.
- Ebel , K.L. (1972) Essentials of Education Measurement , 2nd Erylood : California .
- Filipek, P. J. Accardo, S. Ashwal, G. T. Baranek, E. H. Cook, Jr., G. Dawson, B. (2010) :Standards Subcommittee of the American Academy of Neurology and the Child , Practice parameter: Screening and diagnosis of autism: Report of the Quality
- Ghiselli , E.E. etal, S (1981) Measurement Theory for the Behavioral Science , San Francisco . W.H – Freh man &Company .
- Ghiselli ,E .E (1964), Theory Of Psychological Measurement . New York .Mc Graw –Hill.
- Gilliam, J. (2014) Autism spectrum Disorders International Association for Child and Adolescen Psychology and Allied Professional, TX : Pro-Ed.
- Gilliam, J.E (2013) Gilliam Rating Autism Scale (GARS-3) , Austin , TX : Pro-Ed.
- Harrison, A. (1983) A Language testing Handbook, London, the Mac Millan Press.
- Krus , D. J. ; & Fuller , E.A. (1982) . Illustration of the Theory of True and Error Scores. Computer Based Method of Idealized .
- Magnusoon, D. (1967).Testing Theory, London, Addison – Wesley Publishing Company .
- Melhem ,S. (2000) Measurement and Assessment in Psychology .New vision(1st edn.)Amman:Dar Al Massira for Publishing , Printing and Distribution.
- Nunnally , J. c. (1978) Psychometric Theory , 2nd ed , New York : Mc Graw – Hill
- Samadi, Sayyed Ali & MCconkey Roy, (2014): The Utility of the Gilliam rating scale for identifying Iranian children with Autism.

- Samara , A. (1989) Principles of Measurement and Assessment in Education. Dar Al-Fakar for Publication and Distribution.
- Shiriban, L. (2010) Autism between Science and Fiction .Translated by Fatima Ayyad .Kuwait: Dar Al-Maarefa.
- Theeb, R. (2013) Psychometric Characteristics for Syrian Picture out of Gilliam Scale to Diagnose Autism Disorder. Published Research Paper in Journal of Educational Science.(Vol.40) Appendix (1) College of Educational Science Jordan University.
- Weiner , E.A. & stewart , B.J. (1984) : Assessing Individuals Psychological and Educational Tests and Measurement , New York , Little Brown , Co.

استخدام نظرية المقياس النفسي في صياغة مقياس جيليام GARS-3 لتشخيص اضطراب طيف التوحد

م. عباس فاضل عبد الواحد
جامعة بغداد / كلية الآداب / العراق
abbasfadhil@coart.uobagbaghdad.edu.iq
أ.د. محمد انور محمود
Mohammed.a@ircoedu.uobaghdad.edu.iq
جامعة بغداد / كلية التربية ابن رشد / العراق
أ.د. خالد جمال جاسم
Khalid.jamal@ircoedu.uobagbaghdad.edu.iq
جامعة بغداد / كلية التربية ابن رشد / العراق

التقديم: 2022/6/2 القبول: 2022/7/5 النشر: 2022/12/15
Doi:

الملخص

تهدف هذه الورقة إلى صياغة Gilliam Scale GARS-3 لتشخيص اضطراب طيف التوحد (ADS) وفقًا لاستخدام نظرية المقياس النفسي. ولتحقيق هدف البحث اتبع الباحثون الخطوات العلمية في صياغة المقياس والتحقق من مصداقية الترجمة من خلال التحقق من الدقة والذاتية في نقل جميع أفكار المقياس. ثم قام خبير في اللغة العربية بفحص المقياس المكون من (58) فقرة بستة أبعاد. تم تطبيق المقياس على (500) طالب داخل المراكز الحكومية والخاصة وعلى عينة تم اختيارها حسب نمط الفصل العشوائي تضم أولياء الأمور والمدربين في حكومة بغداد. وبالمثل ، للتحقق من السمات السيكومترية للمقياس ، تم الاعتراف بثلاثة عناصر تحكم لاختيار العناصر. الضوابط هي: التعرف على عامل التحكم من خلال استخدام أسلوب مجموعتين متطرفتين من مجموع النقاط. معيار مصداقية العنصر الذي يعتمد على علاقة ارتباطية إرشادية بين درجة العنصر وإجمالي درجات المقياس باستخدام عامل Pearson الترابطي. نسبة تشبع العناصر حسب عامل التحليل الاستكشافي. تظهر نتائج التحليل صحة جميع العناصر. وقد تم التحقق من مصداقية المقياس بطريقتين هما: المصداقية الصريحة: من خلال عرض تعليمات الاختبار وبنوده على مجموعة من الخبراء في قسم العلوم التربوية والنفسية. مصداقية البناء: من خلال تحليل العامل الاستكشافي: معرفة ثبات المقياس حسب ألفا كرونباخ ، وعامل الثبات (0.941). في ضوء الإجراءات التي اتبعتها الورقة الحالية توصل الباحثون إلى بعض التوصيات والمقترحات.

الكلمات المفتاحية: اضطراب طيف التوحد ، مقياس جيليامGARS-3 ، نظرية المقياس التقليدية.