The adaptation of Gilliam Autism Rating Scale-2 within an Omani context: Some initial findings

Aim: The aim of the present study was to report some initial findings concerning the validity and reliability of the Omani version of the Gilliam Autism Rating Scale-2 (OM-GARS-2).

Sample: The sample of the study included 90 children aged 8-14 years and divided into two groups: Autistic group (n = 45), enrolled in two public centers of autism care in two governorates in Oman, and normal group (n =45), enrolled in two public schools in two governorates in Oman.

Raters sample. A total of 8 teachers (4 males and 4 females) working in these centers rated students on the OM-GARS-2. And 7 teachers (4 males and 3 females) in these schools rated students on the OM-GARS-2. All ratings were performed over a three-week period.

Methodology: To answer the research questions, two types of reliability indictors were computed: (1) test-retest reliability, and (2) internal consistency reliability. Then, Pearson correlation coefficient was computed between students' scores on OM-ABC which is studied by several researchers and the total score and the OM-GARS-2 subscales as well as total score. Finally, a multivariate analysis of variance (MANOVA) was conducted where group (autistic vs. normal) was set as an independent variable (factor) and the OM-GARS-2 subscales and Autism index was set as criterion variable to test the hypothesis that there would be one or more mean differences between groups.

Results: The results of the study showed the OM-GARS-2 had temporal stability and internal consistency reliability. The OM-GARS-2 had criterion (type concurrent) validity and discriminant validity. **Conclusion:** To conclude, the reliability and validity indices of OM-GARS-2 are very

similar to that of the original GARS-2 (Gilliam, 2006) and other studies conducted internationally

Keywords: Gilliam Autism Rating Scale-2, Autism, Omani context.

1. INTRODUCTION

Autism is a disorder among a group of disorders under the umbrella of Autism Spectrum Disorder (ASD) or Pervasive Developmental Disorder (PDD) (American Psychiatry Association, 1994, 2000). This disorder is noticed typically before the age of 3 and it has three defining core features: (a) problems with social interactions, (b) impaired verbal and nonverbal communication, and (c) a pattern of repetitive behavior with narrow, restricted interests (CDC, 2006). In the Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition, Text Revision (DSM-V-TR) these features were reduced to two main characteristics**;** social communication and interaction and restricted, repetitive behavior (Mazefsky, *et al*., 2013). According to the latest report published by the CDC, based upon the data collected by the Autism and Developmental Disabilities Monitoring Network on 8-year-old children living in 11 American provinces in 2010, about 1 in 68 children (or 14.7 per 1,000) were identified with ASD. This new estimate is roughly 30% higher than the estimate for 2008 (1 in 88), roughly 60% higher than the estimate for 2006 (1 in 110), and roughly 120% higher than the estimates for 2002 and 2000 (1 in 150) (CDC, 2010).

The growing rates of autism in recent years have led to considerable interest in its core symptoms and diagnosis. Diagnosis is considered a fundamental and prerequisite step to initiate and introduce special education services for children and adults with autism. Filipek, *et al.*, (2000) proposed that “the diagnosis of autism should include the use of a diagnostic instrument with at least moderate sensitivity and good specificity for autism” (p.475). The authors advocated the Gilliam Autism Rating Scale (GARS), among other measuring tools, as diagnosis tools of autism. A survey published in 2008 found that 40% of school psychologists used the GARS-2 in the majority of their ASD-related assessments (Allen, *et al*., 2008). The GARS-2 is a 42-item informant rating scale designed to assist in the identification and diagnosis of autism and provide information on symptom

severity. The GARS-2 was built based on the definitions of autism that emerged from Diagnostic and Statistical Manual of Mental Disorders- fourth edition, text revision (DSM-IV-TR) (American Psychiatry Association, 2000) and the Autism Society of America. Gilliam (2006) stated that the GARS-2 remains the only normed screening instrument based on these definitions. The GARS-2 was normed using a sample of 1,107 individuals identified as diagnosed with autism and aged between 3 and 22 years.

The GARS-2 can be completed by parents, teachers and/or clinicians. Each of the items is rated on a four-point frequency scale (i.e., 0 = Never Observed, 1 = Seldom Observed, 2 = Sometimes Observed, and 3 = Frequently Observed). The 42 items are grouped to form three 14-item subscales. (a) Social Interaction, (b) Communication, and (c) Stereotyped Behaviors. These three scales are combined to create the overall Autism Index (AI). A parent interview is included which taps into the child's development during the first three years of life, however, item scores from this interview are not factored into the overall AI. The Stereotyped Behaviors subscale focuses on stereotyped behaviors, motility disorders and other unique and atypical behaviors. The Communication subscale contains items that describe verbal and nonverbal behaviors that are indicative of autism. Finally, the Social Interaction subscale defined the individual's ability to relate appropriately to people, events and objects (Gilliam, 2006).

For each of the GARS-2 subscales, the numeric responses from the 14 items are summed into a total raw score. The total raw score is converted to a derived standard score (M = 10, SD = 3). The sum of the standard scores from the three subscales is converted into the overall AI (M = 100, SD = 15), which is standardized to a deviation quotient metric. For non-communicative individuals, the Communication subscale is omitted and the AI is calculated based on the other two subscales. According to the manual, an AI score of 85 or higher indicates a "very likely" probability of autism, scores between 70 and 84 suggest the probability of autism is "possibly," and scores of 69 or below indicate that the probability of autism is "unlikely" (Gilliam, 2006, pp. 31-32).

The GARS-2 manual reported reliability data for both internal consistency and stability (Gilliam, 2006). Internal consistency estimates were .88 for Social Interaction, .86 for Communication, .84 for Stereotyped Behaviors, and .94 for the Autism Index. Corrected test-retest coefficients (1-week interval) based on parent ratings of 37 children with autism were .88 for Social Interaction, .70 for Communication, .90 for Stereotyped Behavior, and .88 for the overall Autism Index. The criterion-related validity was established by computing correlation coefficients between the GARS-2 and the Autism Behavior Checklist subscales (Krug, *et al.*, 1993). The construct-identification validity was established by examining (a) relationships of the GARS-2 subscales scores and age, (b) the internal consistency of the GARS-2 subscales interrelationships, (c) the GARS-2 subscales standard scores and Autism Index correlations, (d) evidence that the GARS-2 has practical value and ability to differentiate autism from other groups (e.g., normal, mental retardation, and multiple disabilities).

Gilliam (2006) discussed several differences between the GARS and the GARS-2 including; (1) the developmental disturbances subscale was revised and converted into an interview form to allow examiners to evaluate the child's development during early childhood. This procedures reduces the time needed for completing the ratings, (2) some items were re-written clearly, (3) demographic characteristics of the normative sample are keyed to the 2000 U.S. census, (4) all new norms were created and the normative sample is more clearly described, (5) the total scores of the GARS-2 were changed from Autism Quotient to Autism Index, (6) guidelines for interpreting subscales scores and the Autism Index were changed, (7) a separate chapter is provided in which discrete target behaviors for each item on the GARS-2 are defined and specific examples are given for applied behavior analysis projects and other research purposes, and (8) a separate booklet "Instructional Objectives for Children Who Have Autism" was developed to assist in the formulation of instructional goals and objectives based on the results of the GARS-2 Tools.

The Omani Context

Oman is situated on the North Eastern corner of the Arabian Peninsula with a population of over 4 million native Arabic speakers. Omani population has no standard scale to diagnosis their children who are at risk to be autistic, they have to go Jordan or Tunisia or Egypt to get a diagnosis. Consequently, there appears to be a need for developing a tool that can help educators and clinician to identify or diagnose the autistic children.

Versions of GARS-2

Li (2005) investigated whether a Chinese version and an English version of the GARS-2 were measuring the same construct. The sample of the study included 20 bilingual Chinese-English speaking parents who had at least one neurotypically developing child ages 2 years through 17 years, and who were immigrants in the United States. Scores on the two versions of the GARS-2 correlated highly and significantly for all subscales and for the Autism Index, suggesting that the two versions are measuring the same construct. The subscales of the Chinese version of the GARS-2 showed acceptable internal consistency. A serious limitation of Li' study is the utilization of a non-clinical sample.

Diken, Diken, Gilliam, Ardic, and Sweeney (2012) conducted a preliminary study to investigate the validity and reliability of a Turkish Version of GARS-2 (TV-GARS-2). Participants included 436 children diagnosed with autism. Data were also collected from individuals diagnosed with intellectual disability, with hearing impairment, and from typically developing children in order to examine discrimination validity of the TV-GARS-2. Coefficient alpha of all subscales and the entire instrument showed acceptable internal consistency. The test re-test reliability coefficients showed acceptable temporal stability. The data provided several indices of TV-GARS-2 construct validity; (1) non-significant correlation with students' chronological age except for Stereotyped Behaviors (r = .15, p <.01), (2) significant interrelationship among TV-GARS-2 subscales (r = .34 to .65, p < .01), (3) item showed acceptable discriminating power, (4) significant corrected correlation coefficients between the Autism index and the three subscales (r = .44 to .60, p < .01), and (5)

The TV-GARS-2 discriminated significantly among four groups of children; intellectual disability group, hearing impairment group, normal development group, and autistic disorder group, suggesting evidence of the TV-GARS-2 discriminant validity.

Al Jabery (2008) conducted a preliminary study to develop a Jordanian Arabic Version of the Gilliam Autism Rating Scale (J-GARS-2). The sample included 100 students aged from three to 13 years and it was divided into two groups (50 students each): students with autism and students with mental retardation. The test re-test reliability coefficients showed acceptable temporal stability. Alpha coefficients and split half reliability showed acceptable internal consistency. The total scores of the J-GARS-2 (Autism Index) correlated highly and significantly with the total score on the Arabic version of the Autism Behavior Checklist (AV-ABC; Krug, Arick, & Almond, 1980;). With the exception of the correlation between J-GARS-2 Stereotyped Behavior subscale and the ABC Sensory subscale, all of the hypothesized correlations were found to be significant and moderate to high in magnitude. The correlations between J-GARS-2 subscales raw scores and age were not significant. All subscales raw scores have a strong correlation with the total score of the J-GARS-2. All subscales raw scores have a strong correlation with each other except for the correlation between the Stereotyped Behavior and the Communication subscales. The GARS-2 discriminated between students with autism group and students with mental retardation on all subscales as of the J-GARS-2 and the Autism index.

Rationale and aims of the study

The growing rates of autism in recent years has led to considerable interest in its core symptoms and diagnosis (CDC, 2010). Furthermore, the challenges faced with differential diagnoses of autism, and the symptomatology of this disorder highlighted the need for assessment tools that contribute to accurate diagnoses. Several measuring tools have been developed and used in Western societies to screen and diagnose autism. However, we know little about the symptoms and diagnosis of autism amongst native Arabs. For example, in Oman, the number of studies conducted

is limited and official statistics are apparently not available. One possible reason that Oman lags behind in autism screening and diagnosis has, in part, been due to the lack of Arabic language measures with acceptable psychometric properties and also to the fact that many Omanis do not have an adequate command of the English language for the use of English language measures. As such, there is a need for a valid and reliable autism diagnostic tool written in Omani and normed on Omani-speaking respondents. Thus, the problem of this study emerged from the need to provide the current tool practices of children and adults with autism in Oman with another valid and reliable instrument to be utilized by professionals to enhance the diagnosis practices. Specifically, the present study reports some initial findings about the psychometric properties of an Omani version of the Gilliam Autism Rating Scale (OM-GARS-2) (Gilliam, 2006).

Questions of the study

The study intended to answer the following questions:

1. What are the correlation coefficients of test-retest and internal consistency reliability of the OM-GARS-2?
2. What are the correlation coefficients between the OM-GARS-2 and the Omani version of Autism Behavior Checklist (OM-ABC) in terms of subscales and total (the Autism Index for OM-GARS-2 and the Total Sum for the OM-ABC) scores?
3. Does the OM-GARS-2 differentiate students labeled with Autism and normal students in terms of each subscale score (stereotyped behaviors, communication, and social interaction) and the Autism Index?

2. METHODS

The researchers used the descriptive approach, by distributing the Checklist to the sample.**Sample**

Autism sample. The autism sample included 45 children (25 males and 20 females) aged between 8 and 14 years (M = 12.3, SD = .61) and enrolled in two public centers of autism care in two governorates in Oman. These centers are supervised by the Ministry of Social Development. Children

in these centers are considered lower-functioning due to significant delays including cognitive, social, and communicative impairments, which hinder them from attending classrooms within their respective schools. Those children are diagnosed to suffer autism based on a clinical diagnosis made by a licensed medical professional and/or psychologist in Oman or in another country.

Raters sample. A total of 8 teachers (4 males and 4 females) working in these centers rated students on the OM-GARS-2. The number of years of teaching experience of those teachers ranged from 2 to 7 years (M = 4.6, SD = .64). Most teacher raters had worked with the student being rated for at least three months prior to the rating. Teaching staff raters were familiar with the general characteristics of autism, as a result of their special education training or work experience. The assessment process was created in order for each student to be rated by the staff member who knew her/him best, while also maximizing the statistical independence of each case being rated. All ratings were performed over a three-week period in the first semester of the school year 2015/2016.

Non-autism sample. The non-autism sample was selected as an available sample from the schools which accepted to involve in the research. It included 45 children (23 males and 22 females) aged 8-14 years (M = 12.6, SD = .47). The children were enrolled in two public schools in two governorates in Oman. A total of 7 teachers (4 males and 3 females) in these schools rated students on the OM-GARS-2. The number of years of teaching experience of those teachers ranged from 2 to 8 years (M = 4.9, SD = .51). Most teachers had worked with the students being rated for at least 4 months prior to the rating. Teaching staff raters were familiar with the general characteristics of ASDs, as a result of their work experience and academic qualifications. The non-autism sample did not suffer any difficulties.

The GARS-2

The GARS-2 is a 42-item behavioral checklist designed to identify persons with autism. The 42 items are grouped to form three 14-item subscales. (a) Social Interaction, (b) Communication, and (c) Stereotyped Behaviors. These three subscales are combined to create the overall Autism Index (AI).

A parent interview is included which taps into the child's development during the first three years of life, but it is not part of the scoring system. The Stereotyped Behaviors subscale focuses on Stereotyped Behaviors, motility disorders and other unique and atypical behaviors.The Communication subscale contains items that describe verbal and nonverbal behaviors that are indicative of autism. The Social Interaction subscale describes the individual's ability to relate appropriately to people, events and objects (Gilliam, 2006). All items of the GARS-2 can be rated on a four-point frequency-based scale that ranged from 0 to 3 (i.e., 0 = Never Observed, 1 = Seldom Observed, 2 = Sometimes Observed, and 3 = Frequently Observed). The GARS-2 can be completed by parents, teachers, and/or clinicians.

Translation of the GARS-2

Two bilingual assistant professors of psychology and special education translated the GARS-2 from English into Omani Arabic using the back-translation method (OM-GARS-2). Two other bilingual assistant professors of psychology and special education, working without referencing to the English version of the GARS-2, independently translated the Arabic version back to English. Finally, one certified translator and a bilingual professor of psychology and special education independently compared the original English version of the GARS-2 with the new English version that was translated back from Arabic, and rated the match between the two versions on a scale of 0 or 1. A score of zero represented no match, whereas a score of 1 represented perfect match. The average percentage of match was 96 % which could be considered highly acceptable (see, Maneesriwongul & Dixon, 2004). Furthermore, interobserver agreement was calculated using SPSS Crosstabs function, which produces a Kappa statistic for level of agreement. According to Cohen (1960), Kappa values lie between -1.00 and 1.00, with zero indicating chance agreement, positive values indicating greater than chance agreement, and negative values indicating less than chance agreement. Landis and Koch (1977) categorized Kappa values from 0.41 to 0.60 as moderate and values above .60 as

substantial levels of agreement. The inter observer agreement Kappa value for the OM-GARS-2 was .75.

Autism Behavior Checklist (ABC)

The ABC was published in 1980 (Krug, *et al*., 1980) and is part of a broader tool, the Autism Screening Instrument for Educational Planning (ASIEP). The ABC is designed to be completed independently by a parent or a teacher familiar with the child who then returns it to a trained professional for scoring and interpretation. Although it is primarily designed to identify children with autism within a population of school-age children with severe disabilities, the ABC has been used with children as young as 3 years of age. The ABC has 57 items and each item is weighted according to the degree to which the characteristic is a symptom of autism [1-is related in a small degree to 4-is related in a strong degree]. For example, "whirls self for long periods of time" receives four points, whereas "does not follow simple commands." receives one point (Krug, *et al*., 1993). The items are grouped into five scales: Sensory, Relating, Body and Object Use, Language, and Social and Self-Help. The Total Score, which is the sum of all items in the five scales, is used as a fundamental indicator of autistic disorder. A cut off score of 67 indicates a high likelihood of autism, a score below 53 indicates a low likelihood of autism, and a score between 53 and 67 indicates the need for more investigations (Lord & Corsello, 2005). Al Hadramy (1995) developed the Arabic version of the ABC in Oman (OM-ABC) using a sample of 114 children aged 10-12 years old. She reported that the ABC has good internal reliability, and good sensitivity and specificity. Overall, studies indicated that the ABC instrument has good psychometric properties to use in the Arabic region.

Procedures

The researchers of this study coordinated the data collection procedures as part of a two-year research project by obtaining necessary official permissions and contacting the autism care centers and public schools. Before starting data collection at the autism care centers, one of the researchers in this study and a research assistant held a meeting with teachers in each center to explain the

purpose of the study and familiarize teachers with the instruments (OM-GARS-2 and OM-ABC). The researcher and the research assistant emphasized the notion that participation in data collection is voluntary and that collected data will be kept confidential and they will be used solely for research purposes. Teachers were encouraged to read the OM-GARS-2 and the OM-ABC carefully before the day of the meeting and they were given the chance to ask questions that were answered by the researchers. This procedure intended to support the internal validity of the study by minimizing raters' bias. Teachers were given two weeks to complete the OM-GARS-2 and OM-ABC and rate their students. Teachers were blind to each other. They were instructed not to discuss students' ratings with each other to keep rating independency. The purpose of administering the OM-ABC was to examine the criterion (type of concurrent) validity of the OM-GARS-2. Two weeks later, the teachers were given and asked again to complete the OM-GARS-2 for their students (including pupils) for purposes of examining the test-retest reliability of the OM-GARS-2. The data collection of the normal sample followed the same procedures of the autism sample. Data collection took place during normal classes at targeted schools. Teachers were given one week to complete the OM-GARS-2 and rate their students. A research assistant individually collected the instruments from both the autism care centers and the public schools and reviewed them to assure their full completion. Only students with complete dataset (98%) were included in the analyses.

3. RESULTS

Question 1. ***What are the correlation coefficients of test-retest and internal consistency reliability of the OM-GARS-2?***

To answer this question, two types of reliability indictors were computed: (1) test-retest reliability, and (2) internal consistency reliability.

For the test re-test reliability, Pearson correlation coefficient was computed for the autistic students' scores on the OM-GARS-2 over the two points of data collection. The results showed that the correlation coefficients were .92 for Stereotype Behavior, .89 for Communication, .91 for Social

Interactions, and .93 for the entire instrument (Autism Index). All correlations were statistically significant at .001. For internal consistency reliability, Cronbach's alpha coefficients were computed for the autistic students to judge the internal consistency of the OM-GARS-2 subscales as well as the entire instrument (Autism Index). Results indicated that Alpha coefficients were .91 for Stereotype Behavior, .90 for Communication, .87 for Social Interactions, and .89 for the entire instrument.

Question 2. ***What are the correlation coefficients between the OM-GARS-2 and the Omani version of Autism Behavior Checklist (OM-ABC) in terms of subscales and total (the Autism Index for OM-GARS-2 and the total sum for the OM-ABC) scores?***

To answer this question, Pearson correlation coefficient was computed between students' scores on OM-ABC and the total score and the OM-GARS-2 subscales as well as total score. Table 1 shows that all Pearson correlation coefficients were statistically significant.

Table 1
Pearson correlations coefficients between OM-GARS-2 subscales raw scores and total score (Autism Index) and the OM-ABC subscales raw scores and total sum

	OM-ABC subscales					
OM-GARS-2 subscales	Sensory	Relating	Body and object use	language	Social and Self-help	ABC Sum
Stereotype Behavior	.52**	.42**	.39**	.33*	.40**	.56**
Communication	.49**	.50**	.48**	.45**	.32*	.60**
Social Interactions	.44**	.43**	.28*	.31*	.29*	.57**
Total (Autism Index)	.47**	.47**	.50**	.41**	.37*	.52**

Note. N = 45. **$p < .01$. *$p < .05$

Question 3. ***Does the OM-GARS-2 differentiate students labeled with Autism and normal students in terms of each subscale score (stereotyped behaviors, communication, and social interaction) and the total score (Autism Index)?***

To answer this question, a multivariate analysis of variance (MANOVA) was conducted where group (autistic vs. normal) was set as an independent variable (factor) and the OM-GARS-2 subscales and Autism index was set as criterion variable to test the hypothesis that there would be one or more mean differences between the groups. A statistically significant MANOVA effect was obtained, Pillais' Trace = .43, F (4, 40) = 16.94, $p < .001$. The multivariate effect size was estimated at .23, which implies that 23% of the variance in the canonically derived dependent variable was accounted for by group factor. A series of one-way ANOVA's on each of the four dependent variables was conducted as a follow-up tests to the MANOVA. Univariate results demonstrated a significant effect ($p < .01$) for Stereotype Behaviour, F (1, 43) = 7.33, partial η^2 = .10, Social Interaction, F (1, 43) = 4.30, partial η^2 = .09, and Communication, F(1, 43) = 11.20 partial η^2 = .13. Partial η^2 can vary in magnitude with $\leq$ .01indicting small effect size, $\geq$.02 to $\leq$ .06 indicating medium effect size, and $\geq$ .07 to $\geq$.14 indicating large effect size. Table 2 show mean differences between the autistic group and the normal group. These differences are presented pictorially in Figure 1.

Table 2
Mean differences between autistic group and normal groups in three subscales of the OM-GARS-2 and the Autism index.

Group/Factor	*Stereotype Behaviour*	*Communication*	*Social Interaction*	*Autism Index*
Autistic	3.3	3.5	3.6	3.5
Normal	2.2	2.4	2.2	2.4
Mean differences*	1.1	1.1	1.4	1.1
Cohen's D	.53	.53	.64	.65

Note. N = 45. Means are scaled out of 4. *All mean differences are significant at .01.

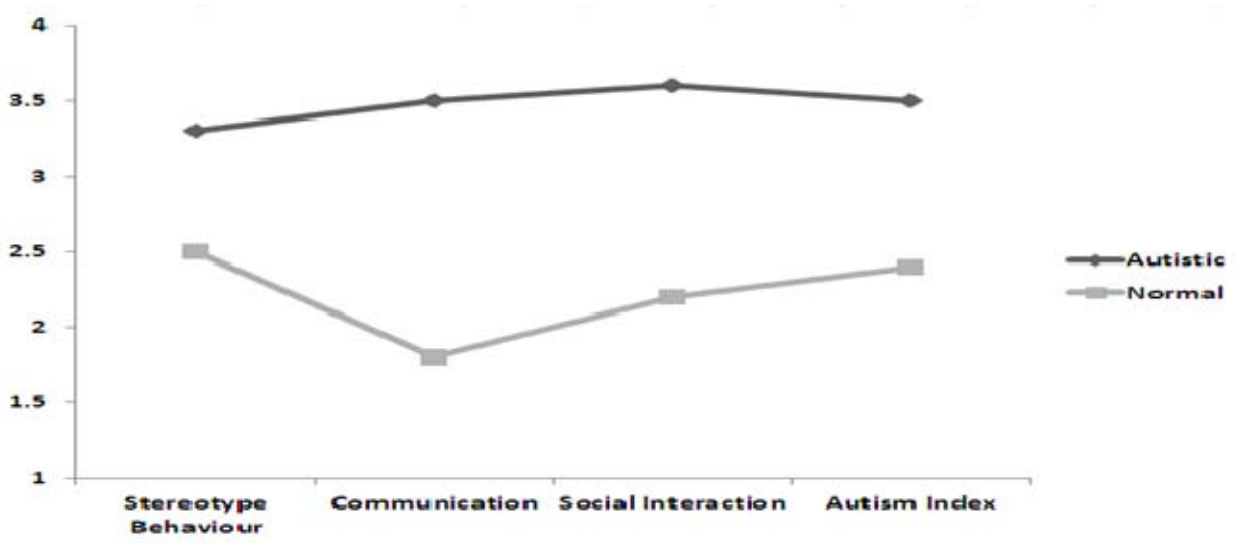

Figure 1. Mean differences between autistic group and normal groups in three subscales of the OM-GARS-2 and the total score (Autism index)

4. DISCUSSION

The present study reports some initial findings concerning the psychometric properties (validity and reliability) of the Omani version of the Gilliam Autism Rating Scale-2 (OM-GARS-2) (2006) within an Omani context. The first step in this study was to translate the GARS-2 from English into Arabic. The goal was to develop an Arabic translated version of the GARS-2 that typically matches the original English version of the scale and that is culturally appropriate for the Omani context. According to the Center for Aging in Diverse Communities, Measurement and Methods Core (2007) “A well-translated survey instrument should have semantic equivalence across languages, conceptual equivalence across cultures, and normative equivalence to the source survey. Semantic equivalence refers to the words and sentence structure in the translated text expressing the same meaning as the source language. Conceptual equivalence is when the concept being measured is the same across groups, although wording to describe it may be different. Normative equivalence describes the ability of the translated text to address social norms that may differ across cultures.” (2007, p.1). The

present study employed the back-translation strategy. A back translation was conducted by an independent translator who has had no previous exposure to the document being translated. Again, the emphasis of the back translation is the conceptual and cultural equivalence (Abd-El-Fattah, *et al.*, 2014). The present study calculated the percentage of agreement of two raters concerning the match between the two versions of the GARS-2; the original English version and the English version that was translated back from Arabic. Although the percentage of agreement was substantially high (96%), the researcher preferred to calculate Kappa statistics for interobserver agreement because percentage agreement does not correct for chance agreement (Abd-El-Fattah, 2013). The kappa statistics was .75 suggested substantial interobserver agreement which implied that the two versions of the GARS-2 were matched. This finding supported the translation of the GARS-2 into Arabic.

The second step was to obtain the psychometric properties for Omani version of the GARS-2 (OM-GARS-2) including validity and reliability indicators to support the entire instrument. The first question of this study concerned the reliability of the OM-GARS2. Reliability indicators were calculated by two methods; the test re-test reliability and the internal consistency reliability. Results indicated strong reliability indicators for the OM-GARS-2 that would support its consistency in measuring the same concept, that is, autism. An examination of the results revealed that the values of the test-retest reliability were high (.92, .89, .91, and .93). One can conclude that the translated version of the GARS-2 has strong temporal stability in measuring the autism disorder. This finding implies that teachers were highly consistent in rating their students during the two times of administration of the OM-GARS-2. This result confirms that teachers understood the OM-GARS-2 items and indicate similar rates. This may imply that teachers as raters are highly educated and well qualified to deal with a rating scale of autism. Moreover, the reliability coefficient of the internal consistency (Cronbach Alpha) were also high to indicate a strong internal consistency. These findings are very similar to that of the original GARS-2 (Gilliam, 2006) and other studies conducted

internationally in Jordan, (Al Jabery, 2008; Alsqour, 2014), Greece (Tafiadis, et *al.,* 2008), Turkey (Diken et *al*., 2008), and China (Li, 2005).

The second question of the present study concerned the validity of the OM-GARS-2. The criterion (concurrent type validity) of the OM-GRAS-2 was established by calculating Pearson correlation coefficient between the OM-GARS-2 and the OM-ABC, hypothesizing that the two instruments are measuring the same construct, that is, autism disorder. The computed correlations coefficients were all statistically significant, suggesting significant correlations between both instruments. One important point to note when examining these correlations is that these correlations were moderate to high in magnitude and that could be attributed either to the small size of sample in which using a bigger sample might improve the correlations, or to raters understanding of the behavioral manifestations of their students' autistic behaviors and their abilities to rate the core deficits as measured by different sample of items (Al Jabery, 2008).

The third question of the present study concerned the discriminant validity of the OM-GARS-2. A multivariate analysis followed by several univariate analyses and post-hoc analyses (Least Significance Difference "LCD") for pairwise comparison showed that the autistic group scored higher than the normal group on the OM-GARS-2 Stereotype Behavior Communication, Social Interaction, and total score. These findings indicated that OM-GARS-2, can be used in differentiating persons with autism from other persons with normal development.

Ethical approval:

We take acceptance from the director of the centres, and the ministry of Education

Limitations and suggestions

Because this research used only 90 subjects (45 autism sample and 45 non-autism sample) which were not enough to report the validity and reliability of the instruments. This should be done in further research with more subjects.

Even though the GARS-2 Omani version has high reliability and validity, it must be noted that this result only comes from are teachers who have expertise and are familiar with observing and evaluating behavioral problems. The reliability and validity of this instrument for the families should be investigated.

CONCLUSION

To conclude, these results offer a valid and reliable tool for autistic children in Oman, to be diagnosis in their country. It help clinician and specialist to have the opportunity to early identification and then, early intervention. Findings of this study also highlights the GARS-2 as a culturally robust scale.

COMPETING INTERESTS

Authors declared that no competing interests exist.

Ethical: NA

Consent: NA

REFERENCES

1- Abd-El-Fattah, S. M. A cross-cultural examination of the Aggression Questionnaire Short form among Egyptian and Omani adolescents. Journal of Personality Assessment. (2013); 95: 539-548.

2- Abd-El-Fattah, S. M., Al Sinani, Y., El Shourbagi, S., Fakhroo, H. A. Using Rasch analysis to examine the dimensionality structure and differential item functioning of the Arabic version of the perceived Physical Ability Scale for Children. Australian Journal of Educational & Developmental Psychology. (2014); 14: 29-44

3- Al Hadramy, M. S. The psychometric properties of the Autism Behavior Checklist (ABC) within an Omani context. Unpublished Master Thesis, Alexandria University, Egypt; (1995).

4- Al Jabery, M. A. The examination of validity and reliability indicators of the Jordanian translated Arabic version of the Gilliam Autism Rating Scale (GARS-2). Unpublished doctoral dissertation, Wayne State University, Michigan; (2008).

5- Allen, R. A., Robins, D. L. & Decker, S. L. Autism spectrum disorders: Neurobiology and current assessment practices. Psychology in the Schools, (2008); 45(10): 905-917

6- Alsqour, I.A. Derivation of performance norms for the Jordanian Arabic version of Gilliam Autism Spectrum Disorder rating scakle for ages 3 to 13 years (Second Edition), Unpublished thesis, Univeristy of Jordon, Amman; (2014).

7- American Psychiatry Association. Diagnostic and statistical manual of mental disorders (4th ed). Washington, DC: American Psychiatry Association; (1994).

8- American Psychiatry Association. Diagnostic and statistical manual of mental disorers (4th ed. Rev). Washington, DC: American Psychiatry Association; (2000).

9- Center for Aging in Diverse Communities. Guidelines for translating surveys in cross cultural research. (2007). Accessed 14 March 2017.
Available: http://www.rcmar.ucla.edu/measurement_surveys.php

10- Centers for Disease Control and Prevention. Parental report of diagnosed autism in children aged 4–17 years—United States, 2003–2004. MMWR. (2006); 55(17): 481-486, Accessed 14 December 2016.
Available: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5517a3.htm

11- Speech & Hearing Associates. (2010). 10 things you need to know about CDC's latest report from the Autism and Developmental Disabilities Monitoring Network. Accessed 29 January 2016.
Available: http://www.cdc.gov/features/dsautismdata/

12- Cohen, J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement. (1960); 20: 37-46.

13- Diken, I. H., Ardic, A., Diken, O, & Gilliam, J. E. Exploring validity and reliability of Turkish version of Gilliam Autism Rating Scale-2. Education and Science. (2012); 37(166): 318-328.

14- Filipek, P. A., Accardo, P. J., Ashwal, S., Baranek, G. T., Cook, E. H., Dawson, G., Volkmar, F. R. Practice parameter: Screening and diagnosis of autism. Report of the quality standards subcommittee of the American Academy of Neurology and the Child Neurology Society. Neurology. (2000); 55(4): 468-479.

15- Gilliam, J. E. Gilliam Autism Rating Scale – Second Edition. Austin, TX: Pro-Ed; (2006).

16- Krug, D. A., Arick, J. R., & Almond, P. J. Autism screening instrument for educational planning. Austin, TX: PRO-ED; (1980).

17- Krug, D. A., Arick, J. R., & Almond, P. J. Autism Screening Instrument for Educational Planning (ASIPE2) (2nd ed.). Austin, TX: PRO-ED; (1993).

18- Landis, J. R., & Koch, G. G. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics. (1977); 33: 363-374.

19- Li, N. (2005). Preliminary validation of the Childhood Autism Rating Scale – Second Edition Questionnaire for Parents or Caregivers (CARS2-QPC) and the Gilliam Autism RatingScale (GARS-2) with a Chinese-speaking population. Unpublished Master Thesis. University of Eastern Kentucky

20- Lord, C., & Corsello, C. (2005). Diagnostic instruments in the autism spectrum disorders. In F.R. Volkmar, R. Paul, A. klin, & D. J., Cohen (Eds.), Handbook of Autism and Pervasive Developmental Disorders (3rd ed., pp. 730–772). New York: John Wiley & Sons, Inc.

21- Maneesriwongul, W., & Dixon, J. K. Instrument translation process: A methods review. Journal of Advanced Nursing Research. (2004); 48: 175-186.

22- Mazefsky, C. A., McPartland, J. C., Gastgeb, R. Z., & Minshew, N. J. Brief report: Comparability of DSM-IV and DSM-5 ASD research samples. Journal of Autism and Developmental Disorders. (2013); 43: 1236-1242.

23-Tafiadis, D., Loli, G., Tsanousa, E, & Tafiadi, M. (2008). The Gilliam Autism Rating Scale (GARS - 2), a pilot study for the Greek autistic population. Poster presented at International Society on Brain and Behavior: 3rd International Congress on Brain and Behavior (28 November – 2 December), Thessaloniki, Greece.