VALIDITY AND RELIABILITY OF TURKISH VERSION OF GILLIAM AUTISM RATING SCALE-2: RESULTS OF PRELIMINARY STUDY

Ibrahim H. Diken
Ozlem Diken
Anadolu University

James E. Gilliam
Austin, Texas

Avsar Ardic
Pamukkale University

Dwight Sweeney
California State University at San Bernardino

The purpose of this preliminary study was to explore the validity and reliability of Turkish Version of the Gilliam Autism Rating Scale-2 (TV-GARS-2). Participants included 436 children diagnosed with autism (331 male and 105 female, mean of ages was 8.01 with SD=3.77). 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. After carrying out Turkish translation procedure, reliability and validity of TV-GARS-2 were explored by conducting a series of analyses. Results yielded that TV-GARS-2 is a reliable and valid assessment tool to be used with individuals with autism in Turkey.

Autism is known to occur around the world regardless of race, culture, and economic class (Trembath, Balandin, & Rossi, 2005). Studies conducted by Baird et al., (2006), Ellefssen et al., (2007), and Gilbert, Cederlund, Lamberg, & Zeijlon, (2006), suggest that approximately one percent of the child population presents with some form of Autistic Spectrum Disorders (ASD). However, there are relatively few tests available outside of English speaking countries for identifying and assessing the disorder. spacing The instruments that are available are often translations of tests that were developed and normed in English speaking countries (e.g., Al Jabery, 2008). The purpose of this study was to translate the Gilliam Autism Rating Scale-Second Edition (GARS-2) into Turkish and conduct psychometric evaluations to determine its efficacy when used in a non-English speaking country.

As a lifelong, trainable, and developmental disorder with an age of onset prior to three years (Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision*; DSM-IV-TR*, 2000; Filipek et al., 1999, 2000; International Classification of Diseases-10; *ICD-10*, 1993; Volkmar et al., 1994), autism is a disorder classified among a group of disorders under Autism Spectrum Disorders (ASD) or Pervasive Developmental Disorder (PDD) (*DSM-IV-TR*, 2000). According to the data from the Center for Disease Control and Prevention (CDC) study in 2000 and 2002 in the US, the prevalence of ASD is 1 in 150 or 6.6 per 1000 of 8-year old children. An average male-to-female ratio of 4.3:1 was also reported (CDC, 2007). Qualitative impairments in social interaction and communication, restricted, repetitive, and stereotyped patterns of behavior, interest, and activities are three common areas that individuals with autisms associated with ASD.

In order to be eligible to receive special education or related treatment services, early identification or diagnosis process is crucial for children with special needs including those with autism. Therefore, screening and assessment procedures in terms of the child find process have to happen as early as possible (Filipek et al., 1999; Stone & DiGeronimo, 2006; Volkmar et al., 1999). Assessment and treatment procedures in ASD are two main areas in which several professional groups and professionals have been interested. For example, American Academy of Pediatrics (The Pediatrician's Role in the Diagnosis and Management of Autistic Spectrum Disorder in Children, 2001; Technical Report: The Pediatrician's Role in the Diagnosis and Management of Autistic Spectrum Disorder in Children, 2001; Identification and Evaluation of Children with Autism Spectrum Disorders: Guidelines for the Clinician Rendering Pediatric Care: Clinical Report by Johnson, Myers, & the Council on Children with Disabilities, 2007), the Child Neurology Society and American Academy of Neurology (The Screening and Diagnosis of Autistic Spectrum Disorders by Filipek et al., 1999, 2000), and American Academy of Child and Adolescent Psychiatry (The Practice Parameters for the Assessment and Treatment of Children, Adolescents, and Adults with Autism and Other Pervasive Developmental Disorders by Volkmar et al., 1999) have been established and provided valuable information regarding assessment processes for individuals with autism (As cited in Al Jabery, 2008). Among these studies, the practice parameter study (Filipek et al., 1999, 2000) stated two levels for the screening and assessment process: Level one and level two. The focus of the level one is the screening process whereas the purpose of level two is the process of diagnose. Johnson, Myers, & the Council on Children with Disabilities (2007) and Filipek et al. (1999, 2000) recommended specific screening tools regarding autism for professionals to use when they have concerns about the individual. Gilliam Autism Rating Scale (GARS) and Gilliam Autism Rating Scale-2 (GARS-2) as level two instruments have been recommended for at risk children 18 months and older. In addition, Coonrod and Stone (2005), and Lord and Corsello (2005) recommended Gilliam Autism Rating Scale (GARS) as non-age specific measures.

Gilliam Autism Rating Scale-2 (GARS-2, 2006) is revised version of Gilliam Autism Rating Scale (1995). GARS-2 was recommended to be used as type two (level two) assessment instruments by Johnson, Myer, and the Council on Children with Disabilities (2007) guidelines.
The GARS and GARS-2 have been used in several studies (e.g., Al Jabery, 2008; Hodge, 2008; Lecavalier, 2005; Mazefsky and Oswald, 2006; Phillips, 2009; Schreck and Mulic, 2000; South et al., 2002; Tafiadis, Loli, Tsanousa, and Tafiadi, 2008). Out of these studies, in two studies GARS-2 was adapted in two different cultures or languages. Al Jabery (2008), for example, in his study, examined reliability and validity of the Jordanian translated Arabic version of the GARS-2 with Jordanian population while Tafiadis, Loli, Tsanousa, and Tafiadi, (2008) studied psychometric characteristics of the GARS-2 with Greek population in Greece. Results of both of these studies showed that adapted GARS-2 had high level of reliability and validity characteristics in both groups of populations. There is currently no standardized or norm-referenced assessment tool in Turkey to screen, diagnose or to be used in the assessment practices of children and adults with autism although reliability and validity of Autism Behavior Checklist-ABC (Krug, Arick, & Almond, 1993) was studied with Turkish sample(Irmak, Sütçü, Aydın, & Sorias, 2007). Therefore, there is an emergent need to have standardized or norm-referenced assessment tool(s) for assessment of individuals with autism in Turkey. This study was designed to provide a valid and reliable either level one or two assessment tool to be utilized by professionals to meet the needs of children and young adults with autism and their families for accurate and appropriate assessment practices in Turkey.

Method

Participants

Participants included 436 children and young adults who had been diagnosed with autism. Out of 436, 331 (75.9%) was male and 105 (24.1 %) was female. Participants' age ranged from 3 to 21 with a mean of 8.01 (SD=3.77).

Data Collection

Participants were selected in a variety of ways. Parents, teachers, psychologists, and other professionals working in private special education and rehabilitation centers in nine cities of Turkey were contacted by the first author and asked to complete the GARS-2 for students with autism. The following criteria were used to be included in the study: (a) having a diagnosis of autism according to psychological testing done in Counseling and Guidance Centers and individuals' health reports, (b) being between ages of 3 and 21, (c) residing in the Turkey. Applying these criteria, 533 participants were reached. However, data collected from 97 of these participants was not included because all items had not re not filled out for these participants. So, the final normative sample included 436 children and young adults with autism.

During the study, additional data were collected on children and young adults who did not have autism. They had diagnoses of intellectual disability and hearing impairment. In addition, data on children and young adults with normal development were collected to establish a control group. These data were also collected from teachers and parents and were used for studies of the discriminating validity of the TV-GARS-2 but were not used in the establishing the norms the instrument. These participants (N=137) were diagnosed with intellectual disability (n= 44) and hearing impairment (n=49). A group with normal development (n= 44) was also selected to serve as a control group.

Measure

Gilliam Autism Rating Scale-2 (GARS-2, 2006). GARS-2 is comprised of three subscales (stereotyped behaviors, communication, and social interaction), with 14 items in each, and a total of 42 items. Items were developed based on the definition of autism adopted by the Autism Society of America and on diagnostic criteria for the autism published in DSM-IV-TR (2000). As indicated in its manual (Gilliam, 2005), GARS-2 can be used for following five purposes: (a) identifying persons who have autism, (b) assessing persons refereed for serious behavior problems, (c) documenting progress in the areas of disturbance as a consequence of special intervention programs, (d) targeting goals for change and intervention on a student's Individualized Education Plan (IEP), and (e) measuring autism in research projects. As a norm-referenced screening instrument, GARS-2 has been used for the assessment of individuals with autism.

GARS-2 was normed on a sample of 1,107 children and young adults aged between 3 and 22 who had been diagnosed as autism. Reliability of validity of the GARS-2 was examined by carrying out a series of psychometric procedures such as content sampling and time sampling for reliability, content-description validity, criterion-related validity, and construct-identification validity for validity. Results of these analyses revealed that GARS-2 a psychometrically sound screening instrument (Gilliam, 2006).

Procedure

Normative Scores. The subscales of TV-GARS-2 are all norm-referenced based on the results from the participants in the normative sample. Raw score means and standard deviations were calculated, and raw scores were then converted to normalized standard scores and percentile ranks. The standard scores are normally distributed and allow an examiner to compare an individual's performance among the three subscales. An individual's score may also be compared to those of the TV-GARS-2 normative sample.

Standard Scores. Standard score norms are expressed as standard deviation units that designate a score's distance from average performance of the normative sample by applying a predetermined mean and standard deviation. For example, the mean and standard deviation for z-score are 0 and 1, respectively; for T-score, they are 50 and 10; and so on. For the TV-GARS-2 subscales, the mean has been set at 10 and the standard deviation at 3 as set in the original GARS-2. Standard scores for the TV-GARS-2 subscale are derived directly from a cumulative frequency table containing the raw scores received by the normative sample. When normative tables are constructed, the raw scores are transformed into the desired derived distribution (i.e., into a distribution with a mean of 10 and a standard deviation of 3). Raw score means and standard deviations were computed for each age and gender. There were minimal differences between participants at different age levels. This is not surprising because the behaviors of autism are not known to differ in terms of age (American Psychiatric Association, 1994). Statistical analyses were undertaken to confirm these observations. Correlations of subscale raw scores with age resulted in correlations of .08 for stereotyped Behaviors (n.s.); .15 for Communication ($p<.01$); and .06 for Social Interaction (n.s.). One-way analysis of variance of TV-GARS-2 subscales scores by gender did not reveal significant differences between males and females on each subscale. The correlation of subscale raw scores with gender resulted in no significances. Correlations of subscale raw score with gender resulted in correlations of .03 for Stereotyped Behaviors (n.s.); .03 for Communication (n.s.); and .03 for Social Interaction (n.s.).

Autism Index. The Autism Index is another type of normalized standard score. This index has a mean of 100 and a standard deviation of 15 (as set in the original GARS-2) and represents the TV-GARS-2's overall assessment of the characteristics of autism manifested by an individual. The Autism Index is derived by summing the standard scores for all subscales of the TV-GARS-2 that were recorded.

Percentile Ranks. Percentile ranks are reported for each of the TV-GARS-2 subscales. Percentile ranks, like standard scores, are derived directly from the raw score distribution of a test. They indicate the

percentage of scores in the normative group that are above or below the score question. Table 1 represents information in converting raw scores to standard scores and percentiles on TV-GARS-2.

Table 1. Converting Raw Scores to Standard Scores and Percentiles

	TV-GARS-2 Subscales			
Standard Score	Stereotyped Behaviors	Communication	Social Interaction	%
1	_	_	_	<1
2	_	_	_	<1
3	_	_	_	<1
4	1-2	1-2	1-3	<1
5	3-5	3-5	4-7	4
6	6-8	6-8	8-10	8
7	9-11	9-11	11-13	19
8	12-14	12-14	14-16	30
9	15-17	15-17	17-20	42
10	18-20	18-20	21-23	60
11	21-23	21-23	24-26	72
12	24-26	24-26	27-30	83
13	27-29	27-29	31-33	89
14	30-32	30-32	34-36	95
15	33-35	33-35	37-39	98
16	36-38	36-38	40-42	>99
17	39-41	39-41	_	>99
18	42	42	_	>99
19	_	_	_	>99
20	_	_	_	>99

Results

Reliability of the TV-GARS-2

Content Sampling. The internal consistency reliability of the items on the TV-GARS-2 was investigated using Cronbach's coefficient alpha. Coefficient alphas were computed for all of the subscales of the TV-GARS-2 using all of the participants. The resulting coefficients for each subscale are Stereotyped Behaviors .82; Communication .81; Social Interaction .87; and for the total test (all 42 items) .91.

Time Sampling. To determine whether the results of the TV-GARS-2 are stable over time, a study was completed in which raters completed the TV-GARS-2 twice, 2 weeks apart, on 35 individuals with autism enrolled in a private special education and rehabilitation center in Eskisehir in Turkey. The raters were parents of the children. The mean age of the children was 6 years (SD=2.8). Twenty-seven of children were male and eight were female. Raw scores for the two testing were converted into standard scores and indexes. The values were then correlated and corrected for restriction in range. The results, reported in the Table 2, provide evidence of the stability of the TV-GARS-2 when used with individuals with autism. The coefficients are all beyond the .01 level of significance and of sufficient magnitude to suggest that the TV-GARS-2 has good test-retest reliability for use as an instrument for identifying persons with autism. These findings demonstrate that the TV-GARS-2 yields results that are stable over time.

Validity of the TV-GARS-2

Translation procedures and face validity. During the Turkish translation process of the GARS-2, six Turkish professionals working in the field of special education provided input about Turkish version of the GARS-2 items. By gathering Turkish translations of items from six professionals, a final version of items was prepared by the first author. After retranslating items to English, they were then translated from English to Turkish by a professional in special education who had excellent English and Turkish skills and was familiar with the characteristics of individuals with autism. Final version of Turkish items of GARS-2 then was tested with a small sample of parents. Fifteen parents filled out the final form and provided input about its face validity.

Table 2. Results of Time Sampling of the TV-GARS-2

TV-GARS-2 Subscales	Time 1		Time 2		Correlations r^c	
	M	SD	M	SD	r	r^c
Stereotyped Behaviors	9	3	9	3	.97*	.96*
Communication	9	2	9	2	.97*	.98*
Social Interaction	9	3	9	3	.96*	.96*
Autism Index	95	15	95	14	.94*	.94*

*p<.01, r^c= coefficent corrected for restricted range

Content-description validity and item analysis. Item-discrimination analysis was conducted to confirm the validity of the test items. Two item discrimination criteria were used test the TV-GARS-2. Using the criteria established by Hammill, Brown, & Bryant (1992), the item-discrimination coefficients had to be statistically significant at or beyond the level .05 level. Second, at least half of the correlation coefficients had to reach or exceed .35 in magnitude. The minimum is large enough to ensure that each item is making a meaningful contribution to the subtest. Conventional item analysis was performed on 301 cases from the sample. These cases were selected because they had complete data; that is, all 42 items of the TV-GARS-2 were completed. In most cases, item analyses are performed for each age interval, but because little relationship exists between age and scores on the TV-GARS-2 subscales, item analyses were not necessary at each age. The results regarding item-discrimination coefficients are reported in Table 3.

Table 3. Item-Discrimination Coefficents for the TV-GARS-2 Subscales

Stereotyped Behavior		Communication		Social Interaction	
Item #	r	Item #	r	Item #	r
1	.43	15	.37	29	.48
2	.25	16	.36	30	.55
3	.35	17	.38	31	.38
4	.32	18	.44	32	.41
5	.33	19	.46	33	.53
6	.37	20	.50	34	.50
7	.40	21	.36	35	.54
8	.45	22	.26	36	.57
9	.39	23	.26	37	.61
10	.38	24	.27	38	.46
11	.38	25	.26	39	.51
12	.37	26	.46	40	.50
13	.49	27	.38	41	.55
14	.42	28	.33	42	.33
Median = .38		Median = .36		Median = .50	

The following median coefficients were obtained; Stereotyped Behavior, .38; Communication, .36; Social Interaction, .50. The median coefficient with sum of all items was .39. The median coefficients for the subscales and sum were statistically significant (p<.01). In addition, they were well beyond the minimum criteria for magnitude and provide ample evidence of content-description validity.

*Construct-identification validity***.** To demonstrate the construct validity of a test, one must delineate as fully as possible the variable (construct) that the test purports to measure. This is done by setting up hypotheses are subjected to scientific investigation, and they are accepted or rejected on the basis of the

results. The following hypotheses (as tested in the original study of the GARS-2) were tested for construct-identification validity of the TV-GARS-2:

1. Because the behaviors measured by the TV-GARS-2 reflect the lifelong nature of autism, TV-GARS-2 scores should not correlate highly with chorological age.
2. Because the TV-GARS-2 subscales are related to each other (i.e., they all contain items that measure some aspect of autism), the subscales of the TV-GARS-2 should be positively related to each other.
3. Because the items within each TV-GARS-2 subscale measure similar traits, the items of a subscale should relate highly with the total score of that subscale.
4. Because the TV-GARS-2 subscales all measure characteristics of autism, they should be positively related to the Autism Index.
5. Because the TV-GARS-2 measures autism, the scores of persons with autism should differ significantly from those of persons who do not have autism.
6. Because the TV-GARS-2 measures autism, the TV-GARS-2 scores should discriminate persons who have autism from those who do not.

Hypothesis 1: Relationship of the TV-GARS-2 subscales to age. To study this hypothesis, the TV-GARS-2 raw scores were correlated with age using data of all participants with autism. The resulting coefficients were found .08 ($p>.01$) for Stereotyped Behaviors, .15 ($p<.01$) for Communication, , .06 ($p>.01$) for Social Interaction, and 14 ($p<.01$) for total (sum of all 42 items.

*Hypothesis 2: Interrelationship among GARS-2 subscales***.** To examine the relationships of the TV-GARS-2 subscales, standard scores of the subscales were correlated for all the participants. Table 4 displays the results of the correlations of the TV-GARS-2 subscales. All of the correlations are moderate to large in magnitude by Hopkins's criteria. Clearly, the items of each subscale measure the same construct (i.e., behavioral characteristics of autism).
Table 4

Table 4. Correlation of TV-GARS-2 Subscale Standard Scores (Decimals Omitted)

	Stereotyped Behaviors	Communication
Communication	34*	–
Social Interaction	65*	53*

*p < .01

Hypothesis 3: Validity of TV- GARS-2 items. The discriminating power of an item, computed by the point-biserial method of item-total correlations, is sometimes referred to as item validity because these coefficients reflect the degree to which the items of a subtest or test are measuring the same constructs. These data can be cited as evidence of a test's reliability because strong item discrimination can only result from strong construct validity. Evidence of the item validities associated with the TV-GARS-2 appears in Table 3. Tests having poor construct-identification validity are unlikely to be composed of items having coefficients of the size reported in this table.

Hypothesis 4: Relationship of subscale standard scores to the Autism Index. To test this hypothesis, the standard scores for all subscales of the participants in the sample were correlated with the Autism Index. Part-whole correlations were computed, meaning that when a subscale was correlated with the Autism Index, its values were excluded from the index before the correlation was computed. In this way, the true correlation of the subscale value to the Autism Index is verified. As can be seen in Table 5, all correlations are large in magnitude and significant ($p < .01$).

Table 5. Correlation of GARS-2 subscales with the Autism Index (Decimals omitted)

	Stereotype Behaviors	Communication	Social Interaction
Autism Index	44*	48*	60*

*p < .01

Hypothesis 5: Differences on TV-GARS-2 standard scores between diagnostic groups. In terms of the TV-GARS-2, individuals who will be tested with this instrument will be those with severe behavioral disorders and developmental disabilities. Therefore, to establish the validity of the TV-GARS-2, a sample of individuals without autism was administered the GARS-2 (N=137). These participants were diagnosed with intellectual disability (n= 44) and hearing impairment (n=49). A group with normal development (n= 44) was also selected to serve as a control group. To test hypothesis that persons from different diagnostic groups will score differently than persons with autism on the TV-GARS-2, One-way analysis of variance procedure was run but homogeneity of variance (Levene Test) is not assumed. Therefore, appropriate nonparametric procedures (Kruskal-Wallis H and Mann-Whitney U test) were run. Results showed that differences between groups were statistically significant. Table 6 reports the mean standard scores for all groups. On each TV-GARS-2 subscale and the Autism Index, the group with autism received significantly higher scores ($p < .01$) than the other diagnostic groups. On the Autism Index, the group with normal development received significantly lower ratings than the other diagnostic groups.
Table 6

Table 6. Mean of standard scores of different diagnostic groups on the TV-GARS-2

		TV-GARS-2 Subscale Standard Scores			
Diagnostic Group	N	Stereotyped Behaviors	Communication	Social Interaction	Autism Index
Intellectual Disability	44	4	5	5	69
Hearing Impairment	49	2	2	1	49
Normal Development	44	1	1	1	46
Autistic Disorder	436	10	10	10	100

Discussion

With the current study, we aimed at providing a valid and reliable assessment tool to be utilized by professionals in Turkey. For this aim, we studied validity and reliability of the Turkish version of GARS-2. Participants in the normative sample covered a wide geographic range coming from 9 cities around Turkey. This diversity adds to the strength of the subscales and provides comparisons for a variety of relevant demographic characteristics. In the study, parents (n=254) completed the subscales more than professionals (n=182). This ratio approximately 3:2 seems appropriate because teachers and other professionals will be the principal raters using the TV-GARS-2. Moreover, the GARS was also found to be useful and recommended as an appropriate instrument for gathering information from parents about their children with autism (Filipek et al., 2000). Since there is no information about prevalence characteristics of individuals with autism in Turkey, we can not decide whether our sample represent the autistic population of Turkey. However, given the male-to-female ratio (3.15:1 in the US) as a reference, the sample used for this study appears representive.

According to Hammill et al., (1992), normative samples should have at least 75 to 100 participants at every age level and at least 750 to 1000 in the total sample. The normative sample for the TV-GARS-2 does not meet this standard for both total sample size and the minimum of 75 participants at all of ages. Results of the current study revealed that there were minimal differences between participants at different age levels. This was not surprising because the behaviors of autism are not known to differ in terms of age (American Psychiatric Association, 1994). Statistical analyses were undertaken to confirm these observations. Correlations of subscale raw scores with age resulted in correlations of .08 for stereotyped Behaviors, .15 for Communication ($p<.01$), and .06 for Social Interaction. Although the correlation for Communication with age is significant, its magnitude is small according to Hopkins' Likert-scale approach in determining the magnitude of coefficients (As cited in Gilliam, 2006). He suggested that coefficients between 0.0 and .09 are very small, coefficients between .1 and .29 are small, coefficients between .3 and .49 are moderate, coefficients between .5 and .69 are large, and coefficients between .7 and .89 are very large. Using Hopkins's guidelines, one must conclude that the relationship between subscale scores and age on the TV-GARS-2 is relatively meaningless. Based on this information, even though age might not be seen as a concern, sample size is a limitation of the current

study. By taking this study as a preliminary study of standardization process of the TV-GARS-2, norming should be done with 750 to 1000 participants.

Results related to internal reliability of total scale and all subscales of the TV-GARS-2 suggest that the items within the subscales are quite consistent and all of the subscales are sufficiently reliable for contributing to important diagnostic decisions. Test-retest data also suggested that TV-GARS-2 yields results that stable over time. So, this information provides confidence in the subscales to professionals when making decisions or interpreting results from the TV-GARS-2.

Results regarding item analyses of TV-GARS-2 showed that the test items provide ample evidence of content-description validity. Although criterion-validity of the TV-GARS-2 could not be checked since there is no standardized assessment tool in Turkey, construct-identification data also provided sufficient evidence for Turkish version of the GARS-2. For example, age was found that it had no meaningful relationship with TV-GARS-2. Using Hopkins's criteria, the correlation coefficients among subscales have been found moderate to large in magnitude. In addition, when examining relationship of subscale standard scores to the Autism Index, large in magnitude and significant correlations have been found. This results supports that Autism Index can be used as the best predictor of the likelihood of the diagnosis of autism.
Regarding discrimination-validity of the TV-GARS-2, results revealed that TV-GARS-2 can be used in differentiating persons with autism from other special needs groups or persons with normal development and the identification of individuals with autism from individuals from other diagnostic groups.

Overall reliability and validity results of TV-GARS-2 are very similar results of other studies (Al Jabery, 2008; Gilliam, 2006) and similar results at some points (Limited information was provided at the short abstract of poster presentation) with the study carried with Greek population (Tafiadis, Loli, Tsanousa, & Tafiadi, 2008). These results support the knowledge base that autism is a universal disorder regardless of culture. Results of this study also provide evidence that GARS-2 is culturally robust scale.

The results of this study should be interpreted with the following main limitation. As we mentioned earlier, according to Hammill et al., (1992), normative samples should have at least 750 to 1000 in the total sample for standardization process of developing a test. Therefore, with 436 sample size this study should be considered as a preliminary study of standardization.

References

Al Jabery, M. A. (2008). *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.
American Academy of Pediatrics. (2001). Committee on Children with Disabilities: the Pediatrician's role in the diagnosis and management of autistic spectrum disorder in children. *Pediatrics, 107* (5), 1221-1226
American Academy of Pediatrics. (2001) Committee on Children with Disabilities: Technical report: the pediatrician's role in the diagnosis and management of autistic spectrum disorder in children. *Pediatrics, 107* (5), 1221-1226.
American Psychiatry Association (1994). *Diagnostic and statistical manual of mental disorders (4th ed.)*. Washington, DC: American Psychiatry Association.
American Psychiatry Association. (2000). *Diagnostic and statistical manual of mental disorders (4th ed. Text Revision)*. Washington, DC: American Psychiatry Association.
Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Chairman, T. (2006) Prevalence of disorders of the autism spectrum in a population cohort children in South Thames: The special needs and autism project (SNAP), *Lancet, 368,* (9531): 210-215.
Center for Disease Control and Prevention (2007). Prevalence of autism spectrum disorders: autism and developmental disabilities monitoring network, 14 Sites, United States, 2002. *MMWR Surveill Summ. 56*(1):12–28
Ellefssen, A., Kampmann, H., Billstedt, E., Gillberg, I.C., & Gillberg, C. (2007). Autism in the Faroe Islands: An epidemiological study, *Journal of Autism and Developmental Disorders,* 33 (3): 703-707.
Filipek, P.A., Accardo, P.J., Ashwal, S., Baranek, G.T.,Cook, E.H. Jr., Dawson, G., et al. (1999).The Screening and diagnosis of autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 29*, 439–484.

Filipek, P.A., Accardo, P.J., Ashwal, S., Baranek, G.T.,Cook, E.H. Jr., Dawson, G., et al. (2000). 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, 55*, 468–479.

Hammill, D. D., Brown, L., & Bryant, B. R. (1992). A consumer's guide to tests in print (2nd ed.). Austin, TX: Pro-Ed.

Hodge, D. (2008). *Children's sleep problems and maternal mental health in mothers of children with and without Autism* [Unpublished doctoral dissertation]. Claremont Graduate University, California.

Irmak, T. Y., Sütçü, S. T., Aydın, A., & Sorias, O. (2007). Otizm davranış listesinin (ABC) geçerlik ve güvenirliğinin incelenmesi [Investigating reliability and validity of Autism Behavior Checklist (ABC)]. *Çocuk ve Ruh Sağlığı Dergisi, 14*(1), 13-23.

Gilbert, C., Cederlund, M., Lamberg, K., & Zeijlon, C. (2006). Brief report. The autism epidemic: The registered prevalence of autism in a Swedish urban area. *Journal of Autism and Developmental Disorders, 36* (6): 429-435.

Gilliam, J. E. (1995). *Gilliam Autism Rating Scale (GARS).* Austin, TX: Pro-Ed.

Gilliam, J.E., (2006). Gilliam Autism Rating Scale: Second Edition. Austin, TX.: PRO-ED.

Johnson, C.P., Myers, S.M., & Council on Children With Disabilities. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics,120 (5)*, 1183–1215. Retrieved November 16, 2009, from www.aap.org/pressroom/AutismID.pdf

Lecavalier, L. (2005). An evaluation of the Gilliam Autism Rating Scale. *Journal of Autism and Developmental Disorders, 35* (6), 795— 805.

Mazefsky, C., & Oswald, D. (2006). The discriminative ability and diagnostic utility of the ADOS-G, ADI-R, and GARS for children in a clinical setting. *Autism, 10*(6), 533-549.

Phillips, N. (2009). *A survey of the work and interactions of school psychologists, school neuropsychologists and clinical psychologists in schools today.* Unpublished doctoral dissertation, Claremont Graduate University, California.

Schreck, K. A., and Mulick, J. A. (2000). Parental report of sleep problems in children with Autism. *Journal of Autism and Developmental Disorders, 30* (2), 127-135.

South, M., Williams, B. J., McMahon, W. M., Owley, T., Filipek, P. A., Shernoff, E., Corsello, C., Lainhart, J. E., Landa, R., & Ozonoff, S (2002). Utility of the Gilliam Autism Rating Scale in research and clinical populations. *Journal of Autism and Developmental Disorders, 32* (6), 593–599.

Stone, W.L., & DiGeronimo, T.F. (2006). *Does my child have autism? : a parent's guide to early detection and intervention in autism spectrum disorders*. Jossey-Bass: San Francisco.

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 Behaviour: 3rd International Congress on Brain and Behaviour, Thessaloniki, Greece. 28 November – 2 December 2007.

Trembath, D. Balandin, S., & Rossi, C. (2005). Cross-cultural practice and autism. *Journal of Intellectual & Developmental Disability, 30*, 240-242.

Volkmar, F.R., Cook, E.H., Pomeroy, J., Realmuto, G., & Tanguay, P. (1999). Practice parameters for the assessment and treatment of children, adolescents, and adults with autism and other pervasive developmental disorders. American Academy of Child and Adolescent Psychiatry Working Group on Quality Issues. *Journal of the American Academy of Child and Adolescent Psychiatry*, *38 (12),* 32S–545.

Volkmar, F.R., Klin, A., Siegal, B., Szatmari, P., Lord, C., Campbell, M., et al. (1994). Field trial for autism in DSM-IV. *American Journal of Psychiatry, 151,* 1361-1367.

World Health Organization (1993). *The ICD-10 classification of mental and behavioral disorders: Clinical descriptions and diagnostic guidelines.* Geneva, Switzerland: World Health Organization.