A Reliability and Validity Study of a Rating Scale for Attention Deficit Hyperactivity Disorder in Egyptian Children and Adolescents

M. H. El-Defrawi, R. Mahfouz and L. Ragab

The Conners Parent-Teacher Rating Scale has been used widely in assessment and treatment studies of Attention Deficit Hyperactivity Disorder (ADHD) and proved to be of high reliability and validity. This study reports on the reliability and validity of the Arabic version of Conners Rating Scale as used by parents, teachers and other raters in three different groups of Egyptian children and adolescents. A healthy group of normal preschoolers, a medically referred group of children, and a group of children and adolescents with the diagnosis of ADHD.

Test-retest and inter-rater reliability were analyzed. The high degrees of reliability for the normal preschoolers (r=0.96) and the ADHD group (r=0.64 and 0.68) in contrast to the medically referred group (r=0.53) are discussed. It is emphasized that such a scale should be under no circumstances used as a means for diagnosing ADHD in children, rather it should be viewed as providing additional data that may facilitate the interpretation of other clinical material and assist in the assessment of treatment efficacy.

(Egypt.J. Psychiat.,1992,15:38-44).

Introduction

Attention Deficit Hyperactivity Disorder (ADHD) is a well recognized disorder with three key clinical signs: short attention span, motor activity and impulsivity. Estimates of the prevalence of ADHD in the general population have reached up to 10% and between 3 and 5% of the primary school population (Wender, 1971). Because of the nature and the situational variability of the ADHD, i.e., being worse in group situations and in settings that make demands on the child which provide little structure or clear guidelines for expected behaviour it is helpful to use a tool to evaluate the presence of a disorder, its severity, and even its response to treatment. Rating scales have proved to be quite helpful in evaluating symptomatology in childhood psychiatric and behavioural disorders as exemplified by the broad use for assessing attention deficit disorder of the Conners Behaviour Check list, an instrument with considerable degree of validity and reliability. Since Conners (1969) had introduced the parent-teacher rating scale for evaluation of hyperkinetic children it has been used successfully in diagnostic and treatment studies of the behavioural problems in children (Conners, 1973). This is particularly so for the busy clinician who needs to know about the child performance and behaviour at home or in school . This is true for the ADHD for which norms are available and drug effects can be detected reliably with this scale (Gittelman-Klein, 1980). In Egypt, disruptive behavioural problems have reached 25% (including ADHD, 20%) of the psychiatric presentations to a child

Mohamed H. El-Defrawi, M. D., Lecturer of Psychiatry, Suez Canal University.

Refaat Mahfouz, M. D., Professor of Psychiatry, El- Menyia University.

Lamis Ragab, M. D., Assistant Professor of Pediatrics, Cairo University.

psychiatry clinic within the school health system. (El-Defrawi and Mahfouz, 1990). We are not aware of any published report on the reliability and validity of any diagnostic instrument carried out with Egyptian children and adolescents. In the present study, the authors have attempted to test the reliability and validity of an instrument modified upon the diagnostic criteria of ADHD of DSM-III-R that could be applied to children and adolescents when administered by parents or by others with detailed knowledge of their behaviour. The following is a report on the reliability and validity of the Arabic version of a rating scale for ADHD.

Subjects And Method

The Conners parent-teacher rating scale is a very simple 10-item scale that is usually scored by parents and/or teachers and seems to be an excellent screening instrument in disruptive behaviour disorders in children and adolescents(see the Appendix). A total score is derived from the scale and the cutoff score of 15 has been established as the point that confirms the presence of ADHD since it is very much above the scores received by normal children (Sprague et al. 1974).

In order to facilitate the practical use of this scale, it was translated to Arabic and was given to referees to comment on the adequacy and fluidity of the items compared with the original version. In addition, the scale in its Arabic translation was used in a pilot study to test and detect any inadequacies requiring modification, rewording or reformulation of the items.

The study consisted of three different parts. The first part consisted of using the scale to evaluate and rate 30 preschoolers (age 5-6 years) by two independent teachers in a public nursery school. Teachers were blind as to the aim of the study. They were asked to observe each child separately for ten minutes on three different occasions and then rate each child separately. The ratings of both teachers were analyzed as an inter-rater reliability study.

The second part consisted of rating 37 child and adolescents diagnosed as ADHD. There were two sets of ratings, the first set of ratings was done by parents (usually the father and mother) or a parent and an attendant (for single parents, n=5). thirty seven pairs of ratings constituted the inter-rater reliability study for the ADHD group. Test-retest reliability was done for only 20 of the ADHD children and adolescents of the 37 with one or two weeks between the first and second ratings by the same rater. All children in this group fulfilled DSM- III-R criteria for the diagnosis of ADHD (American Psychiatric Association, 1987). Children and adolescents were referred to a child psychiatric clinic within the school health system in Suez City (El-Defrawi and Mahfouz, 1990).

The third part of the inter-rater reliability study was carried out on 30 children and adolescents attending a pediatric thalassemia clinic in a university hospital. Ages ranged from 4 to 5 years. Raters were blind to the aim of the study. They were asked to rate each child or adolescent separately. The first rater was a pediatric resident and the second was one of the parents who accompanied the patient.

Inter-rater and test-retest reliability were assessed using the Pearson's correlation coefficient.

Results

Demographic variables are presented in table (I). The mean age of the children was 7 years and 8 months. Approximately 90% of the children were prepubertal, with more boys than girls. Table (2) shows the comparison of mean raw scores in the three studied groups. Table (3) shows agreement and disagreement, i.e. inter-rater and test-retest for the three groups of children.

Based on the ratings of the normal

Table 1

Demographic Variables of Children and Adolescents

Variable	Subjects		
	Normal	Hyperactive	Thalassemia
Sex	Number	Number	Number
Female	16	9	13
Male	14	28	17
Age			
< 7 yrs.	30	17	13
7 - 12 yrs.	0	16	12
13 - 17 yrs.	0	4	5

Mean age:
Normal group 6 years (range 5-6)
Hyperactive group 8 years (range 3-14)
Thalassemia group 9 years. (range 5-15)

preschoolers by teachers in the first part of the study, 3 children were rated as hyperactive by the first and second teachers and a fourth child by the second teacher alone. Means (±SD) of the total scores of teacher 1 and teacher 2 are: 4.8 (±7.1) and 5.2 (±7.4) respectively. Notably the mean score on the scale for the normal preschoolers did not differ significantly from the mean score for the thalassemia major group. Inter-rater agreement reached near excellent ($r^{\wedge}$ = 0.96). The correlation coefficient in a similar sample of the general population is represented by $r^{\wedge}$ which is equal to 0.96.

Table 2

Comparison of Mean Raw Scores in the Three Studied Groups

A.	Normal Preschoolers (N=30)	5.0 ±	7.2
B.	Hyperactive group (N=37)	22.1 ±	4.9
C.	Thalassemia group (N=30)	8.1 ±	5.5

A & B, $p<0.01$
A & C, NS
A & C, $p<0.01$

The second part of the study included a comparison of agreement and disagreement between raters (inter-rater) for 20 ADHD children. The means of the total score of hyperactive children rated by mother and father (or teacher when only one parent is available) are 19.5 (±6.2) and 19.7 (± 5.8) respectively. Inter-rater reliability was 0.67. This is still considered to be in the range of a good agreement. ($r^{\wedge}$= 0.64). Test-retest reliability of ratings for 37 children with ADHD resulted in means (±SD) of 23.0 (±5.2) and 21.2 (±4.6) for the first and second assessments respectively. Test-retest reliability showed correlation coefficient of 0.69 ($r^{\wedge}$ = 0.68). This is still considered a good agreement.

The third part of this study consisted of assessment of medically referred group of children, with thalassemia major (n=30). Inter-rater reliability was the lowest $r^{\wedge}$= 0.55 ($r^{\wedge}$= 0.53). Means (±SD) of the total scores were 6.9 (±5.1) and 9.4 (±6.0) for the first rater (pediatric resident) and the second rater (father or mother) respectively.

The use of Conners scale for ADHD picked up three hyperactive children within the group of normal preschoolers There was a disagreement between both teachers on a fourth preschooler whether he was hyperactive or not . However, this might be the incidence of 10% which represents the normal distribution of ADHD in the general population (Wender, 1971). This finding suggests that this scale may also be useful as a screening instrument for identifying children in non-psychiatric

Table 3

Agreement and Disagreement Between Raters: Inter-retar and Test-retest (cut-off score ≥ 15)

Assessed Group	Teacher A		Teacher B		Pearson's Correlation Coefficient
	Nor.	Hyper.	Nor.	Hyper.	
Normal Preschoolers (N=30) (inter-rater)	27	3	26	4	
Mean	4.8		5.2		0.96
S.D.	7.1		7.4		r^0.96
	Rater 1		Rater 2		
ADHD (N=20) (inter-rater)	7	13	5	15	
Mean	19.5		19.7		0.67
S.D.	6.2		5.8		r^0.64
ADHD (N=37)	Time 1		Time 2		
(test-retest)	3	34	1	36	
Mean	23.0		21.2		0.69
S.D.	5.2		4.6		r^ 68
	Rater 1		Rater 2		
Thalasseemia group (N=30) (inter-rater)	29	1	25	5	
Mean	6.9		9.4		0.55
S . D .	5.1		6.0		r^ 0.53

Nor. = Normal
Hyper = Hyperactive

settings that are at high risk for ADHD. Of course, conclusive evidence for this would require a detailed evaluation of positive children as well as a sample from the population with scores in the normal range to determine the validity of the scores.

Discussion

Over the last decade, it has become increasingly evident that hyperactivity and attention problems are significant contributors to serious psychiatric disturbances in childhood and adolescence (Szatmari et al. 1989). In fact, some data suggest that the prevalence of this disorder may be increasing (Costello, 1989). For those interested in rating scales of childhood psychiatric disorders it is reassuring to find agreement between raters and across time on the presence or absence of the signs of ADHD in normal, hyperactive and medically referred groups. It is important, however, to understand the differences in the assessment process in the three sets of the study in order to analyze the data appropriately. Young et al. (1987), in an excellent review of research on clinical diagnostic instruments, pointed out that clinicians training and working together on the use of a structured assessment will commonly achieve agreement in the range of 0.8 (Pearson correlation coefficient). Agreement between parents regarding ratings of behavioural and emotional items, or between other pairs playing a generally similar role in relation to the child (e. g. a pair of teachers), reach a range around 0.6. Different types of informants (e.g., parent and teacher) can be expected to demonstrate an agreement level no better than approximately 0.3 (Achenbach et al. 1987).

In our inter-rater comparison, or test-retest reliability, the correlation coefficient ranged from 0.96 for the normal preschoolers to 0.53 for the

group with thalassemia major indicating near excellent to good agreement. The finding that the inter-rater agreement was lowest for the medical group (r^=0.53) might be explained by the association of generalized anxiety, hyperactivity and oppositional behaviour in chronic medical conditions (Galler and Ramsey, 1989). One can speculate that when the children are confronted repeatedly with the clinic setting and frequent medical consultations they exhibited more disruptive behaviour (Shaw and Emery, 1989) as a defence against anxiety. In general, parents tended to overscore and the pediatric residents to underscore the behavioural problems. This might be a reflection of the psychopathology of parents or their stressful distorted perception of the medical condition of their children.

The second part of this study has shown that although the inter-rater and test-retest of the ADHD were good (0.64 and 0.68 respectively), however, the severity or the presence of psychopathology does not always correlate directly with high agreement on behaviour between raters. It was, therefore, reassuring to find that disagreement between parents (25-35% of the ADHD group was missed) was higher than disagreement between teachers in the normal preschoolers (10% have received the ADHD diagnosis). In this respect, Verhulst et al. 1989 have pointed out to the mother-father disagreement on the information diagnostic of ADHD. However, ADHD is a clinical diagnosis and needs multiple sources of information. Thus, Conners scale appears to distinguish, with good and accepted precision, between normal and hyperactive disturbed children. This Conners ADHD scale, when applied to children diagnosed as having ADHD, appeared to very effectively differentiate them from children with no psychiatric complaints and/or problems (normal preschoolers) and from children with chronic blood disease (Thalassemia major). These preliminary results suggest that the rating scale may have face validity for distinguishing ADHD from non-psychiatric children and adolescents. The scale requires one or two minutes to be filled out and does not need sophisticated skills. It seems reasonable to assume that obtaining good results from all these tests would support reliability and validity of the scale. With the assumption that usually normal children obtain an average score on this scale in the range of 5.0 (±7.2), then probably a score greater than this average by 2 SD basis (i.e., > 2 Standard deviations above the mean) would indicate hyperactivity and attention problems.

However, we think that, on the basis of our clinical findings, the score of 15 is low, especially for young children (6 to 8 years old) and that in spite of a statistical distinction between normal children and children referred for ADHD, on this scale, it may lead to identification of false positives (Gittelman-Klein, 1980). Taken together, the present, preliminary results suggest that the instrument may be a sensitive and specific way for identifying children suffering from ADHD. This conclusion must, however, be tempered by the limitations of the design of the study.

Further studies must now be conducted on this instrument including more work on different subgroups of children in the hope of finding the particular strengths and weaknesses of the different sources of information about the child. Additional studies will be required to assess the validity, sensitivity, and specificity of this instrument in different situations with different groups of children and adolescents.

References

Achenbach, T. M., Mc Conaughy, S. H.and Howell, C. T. (1987) Child/adolescent behavioural and emotional problems: Implications of cross-informant correlations for situational specificity. *Amer. Psychol.*, **191**: 213-232.

Conners, C. K. (1969) A teacher rating scale for use in drug studies with children. *Amer. J. Psychiat.*, **126:** 152-156.

---------- (1973) Rating scale for use in drug studies with children. *Psychopharm. Bull.* Washington, D. C. Department of Health, Education and Welfare.

Costello, E. (1989) Child psychiatric disorders and their correlates in a primary care pediatric sample. *J. Amer. Acad. Child. Adol. Psychiat.*, **28:** 851-855.

El-Defrawi, M. H. and Mahfouz, R. (1990) Psychiatric problems of school-age children and adolescents in a child psychiatric clinic in Suez. Paper read at the *Third International Egyptian Congress of Psychiatry.* 31st May - 2nd June, 1990, Cairo-Egypt.

Galler, J.R. and Ramsey, F. (1989) Follow-up study of the influence of early malnutrition on development: Behaviour at home and at school. *J. Amer. Acad. Child Adol. Psychiat.*, **28:** 251-261.

Gitteman-Klein, R. (1980) Diagnosis and drug treatment of childhood disorders. In: Klein, D.F., Gittelman, R., Quitkin, F.and Rifkin, A. (eds.), *Diagnosis and Drug treatment of Psychiatric Disorders.* 2 nd ed. Baltimore: Williams & Wilkins.

Shaw, D. and Emery, R. (1988) Chronic family adversity and school-age children's adjustment. *J. Amer. Acad. child Adol. Psychiat.*, **27:** 200-206.

Sprague, R. L., Christensen, D. E. and Werry, J.S. (1974) Experimental psychology and stimulant drugs. In Conners, C. K. (ed.), *Clinical Use of Stimulant Drugs in Children.* Amsterdam: Excerpta Medica.

Szatmar, P., Boyle, M. and Offord, D. (1989) ADDH and conduct disorder: Degree of diagnostic overlap and differences among correlates. *J. Amer Acad child. adol. Psychiat.*, 28: 865-872.

Verhulst, F.C., Achenbach, T.M. and Akkerhuis, G.W. (1989) Problems reported for clinically referred American and Dutch children. *J. Amer. Acad. Child. Adol. Psychiat.*, **28:** 516-524.

Wender, P. H. (1971) *Minimal Brain Dysfunction in Children.* New York: Wiley.

Young, J.G., O'Brien, J.D., Gutterman E.M. and Coben, P. (1987) Research on the clinical interview. *J. Amer. Acad. Child. Adol. Psychiat.*, **26:** 613-620.

Appendix
Parent-Teacher Rating Scale for Attention Deficit Hyperactivity Disorder

Not at all = 0
Just a little= 1
Pretty much = 2
Very much= 3

1- Restless (Overactive)
2- Excitable, impulsive
3- Disturbs other children
4- Fails to finish things he starts (Short attention span)
5- Fidgeting
6- Inattentive, distractible
7- Demands must be met immediately
8- Cries
9- Mood changes quickly
10-Temper outbursts (explosive and unpredictable behaviour).

Une Étude de fiabilité et de Validité d'une Échelle d'Évaluation du Trouble d'Hyperactivité avec Déficit de l'Attention Chez les Enfants et les Adolescents Égyptiens

Cette étude éxamine la fiabilité et la validité de la version arabe de l'échelle de Conners appliquée par les parents, instituteurs et autre évaluateurs sur 3 groupes differents d'enfants et d'adolescents Egyptiens. Un premier groupe d'enfants d'âge préscholaire normaux, un second groupe d'enfants consultant une unité médicale et un troisième groupe d'enfants et d'adolescent ayant le diagnostique de trouble d'hyperactivité avec déficit de l'attention. Il nous a paru que cette échelle ne devrait en aucun cas être utilisée comme moyen diagnostique, plutôt elle pourrait fournir des données supplementaires éclairant ainsi l'interpretation du materiel clinique et fournir une mesure de l'efficacité du traitement.

دراسة فعالية ومصداقية مقياس لاضطراب فرط النشاط وقصور الانتباه فى الأطفال والمراهقين المصريين

تتناول هذه الدراسة استخدام النسخة العربية لمقياس كونرز (نسخة الوالدين والمعلم) فى تقويم ثلاثة مجموعات مختلفة من الأطفال والمراهقين لتقدير فعالية ومصداقية قياس اضطراب فرط النشاط وقصور الإنتباه، وقد تكونت المجموعة الأولى من أطفال طبيعين فى سن ما قبل المدرسة، والثانية من أطفال ومراهقين فى عيادة طبية، والمجموعة الثالثة بين أطفال ومراهقين يعانون إضطراب فرط النشاط وقصور الإنتباه.

وقد أشارت النتائج إلى إرتفاع مصداقية إعادة تطبيق المقياس والإتفاق بين المحكمين بالنسبة للأطفال الطبيعين (/ = ٩٦٪)، وكذلك فى مجموعة إضطراب فرط النشاط (/ = ٦٤،. ٦٨،.) بعكس مجموعة العيادة الطبية (/ = ٥٣،.)، وتوضح النتائج والمناقشة فائدة هذا المقياس الذى يجب ألا يستخدم كأداة للتشخيص بمفرده بل فى التقويم المتكامل وتقييم فعالية العلاج.