NEUROLOGÍA

SOCIEDAD ESPAÑOLA
DE NEUROLOGIA

NEUROLOGÍA

PUBLICACIÓN
OFICIAL DE
LA SOCIEDAD
ESPAÑOLA DE
NEUROLOGIA

Volume 32
Number 3
April 2017

ORIGINAL ARTICLE

Validation of a Spanish-language version of the ADHD Rating Scale IV in a Spanish sample[☆],[☆☆]

M. Vallejo-Valdivielso [a],[*], C.A. Soutullo [a],[b],[c], P. de Castro-Manglano [a],[b],[c], J.J. Marín-Méndez [a],[c], A. Díez-Suárez [a],[c]

[a] Unidad de Psiquiatría Infantil y Adolescente, Departamento de Psiquiatría y Psicología Médica, Clínica Universidad de Navarra, Pamplona, Navarra, Spain
[b] Unidad de Psiquiatría Infantil y Adolescente, Departamento de Psiquiatría y Psicología Médica, Clínica Universidad de Navarra, Madrid, Spain
[c] Instituto de Investigación Sanitaria de Navarra (IDISNA), Navarra, Pamplona, Spain

Abstract

Objectives: The purpose of this study is to validate a Spanish-language version of the 18-item ADHD Rating Scale-IV (ADHD-RS-IV.es) in a Spanish sample.

Methods: From a total sample of 652 children and adolescents aged 6 to 17 years (mean age was 11.14 ± 3.27), we included 518 who met the DSM-IV-TR criteria for ADHD and 134 healthy controls. To evaluate the factorial structure, validity, and reliability of the scale, we performed a confirmatory factor analysis (CFA) using structural equation modelling on a polychoric correlation matrix and maximum likelihood estimation. The scale's discriminant validity and predictive value were estimated using ROC (receiver operating characteristics) curve analysis.

Results: Both the full scale and the subscales of the Spanish-language version of the ADHD-RS-IV showed good internal consistency. Cronbach's alpha was 0.94 for the full scale and $\geq$ 0.90 for the subscales, and ordinal alpha was 0.95 and $\geq$ 0.90, respectively. CFA showed that a two-factor model (inattention and hyperactivity/impulsivity) provided the best fit for the data. ADHD-RS-IV.es offered good discriminant ability to distinguish between patients with ADHD and controls (AUC = 0.97).

Conclusions: The two-factor structure of the Spanish-language version of the ADHD-RS-IV (ADHD-RS-IV.es) is consistent with those of the DSM-IV-TR and DSM-5 as well as with the model proposed by the author of the original scale. Furthermore, it has good discriminant ability. ADHD-RS-IV.es is therefore a valid and reliable tool for determining presence and severity of ADHD symptoms in the Spanish population.

Validación de la versión en español de la escala de evaluación del trastorno por déficit de atención e hiperactividad (ADHD-RS-IV.es) en una muestra española

Resumen
Objetivos: El objetivo de este estudio es validar la versión en castellano del ADHD-RS-IV (ADHD-RS-IV.es) en una muestra española.
Métodos: A partir de una muestra total de 652 niños y adolescentes (rango edad: 6-17 años; media [DE] = 11,14 [3,27] años), se incluyó a 518 pacientes con TDAH (criterios DSM-IV-TR) y a 134 controles sanos. Para la evaluación de la estructura factorial, la validez y la fiabilidad de la escala se realizó un análisis factorial confirmatorio (CFA) mediante *structural equation modelling* en una matriz de correlaciones policóricas, y usando el método de máxima verosimilitud para la estimación. Se calcularon la validez discriminante y su valuar predictivo mediante curvas *receiver operating characteristics*.
Resultados: La escala en castellano mostró una consistencia interna elevada, tanto para la escala total como para sus subescalas. El coeficiente alfa de Cronbach era 0,94 para la escala total y ≥ 0,90 para las subescalas. Los valores alfa ordinales eran 0,95 para la escala total y ≥ 0,90 para las subescalas. El análisis CFA mostró un modelo de 2 factores (inatención e hiperactividad/impulsividad) intercorrelacionados. La escala ofrece buen poder discriminante (AUC = 0,97).
Conclusiones: La versión española del ADHD-RS-IV (ADHD-RS-IV.es) mostró una estructura bifactorial consistente con los modelos del DSM-IV-TR y DSM-5, y con el modelo propuesto por el autor de la escala original. Además, posee un alto poder discriminante, lo que lo convierte en un instrumento válido y fiable para medir la presencia y severidad de síntomas de TDAH en la población española.

Introduction

Attention-deficit/hyperactivity disorder (ADHD) is a neurodevelopmental disorder characterised by persistent symptoms of inattention and/or hyperactivity-impulsivity not corresponding to the patient's neurodevelopmental stage.[1] These 2 dimensions are the result of factor analysis of the 18 diagnostic criteria for ADHD listed in the DSM-IV-TR[2] and the DSM-5.[1] A thorough examination of ADHD symptoms should include a clinical interview with the patient and his/her parents and the collection of data on the impact of the symptoms on the patient's overall performance, both at home and at school.[3–8]

Various questionnaires facilitate the collection of data on symptom severity and frequency from teachers and parents, following the criteria and factor structure of the DSM-IV-TR. These tools include the ADHD Rating Scale IV (ADHD-RS-IV),[9] the ADHD Symptoms Rating Scale,[10] the Conners Comprehensive Behaviour Rating Scale,[11] and the SNAP-IV Rating Scale.[12] In response to the changes introduced in the DSM-5,[1] DuPaul et al.[3] recently developed the ADHD Rating Scale-5, supporting the validity of the two-factor structure described in the DSM-IV-TR[2] and DSM-5.[1] The clinical usefulness and the consistent psychometric properties of the ADHD-RS-IV[13] have been replicated in various North American,[14] multi-national,[15] and European studies.[16]

In Spain, Servera and Cardó[17] published normalised data for the ADHD-RS-IV, based on a sample of school-age children (6-12 years). However, the Spanish-language version of the study has not been validated, to our knowledge. The objectives of this study are firstly, to describe the frequency of ADHD symptoms in a sample of children with ADHD and a sample of healthy controls, and secondly, to validate the Spanish-language version of the ADHD-RS-IV (ADHD-RS-IV.es).

Subjects, material, and methods

Subjects

We performed a naturalistic, observational study including 652 participants aged between 6 and 17 years. Participants were divided into two different groups. The patient group included 518 patients with previously untreated ADHD (all subtypes) aged between 6 and 17 years at the time of diagnosis. Patients were selected from among patients diagnosed with ADHD and attended at the child and adolescent psychiatry unit at Clínica Universidad de Navarra between January 2000 and April 2015. The control group included 134 children and adolescents not diagnosed or suspected of having ADHD, aged between 6 and 17 years. Controls were selected from among patients attending the Clínica Universidad de Navarra paediatrics department for routine check-ups in 2014.

Clinical assessment of patients with ADHD

ADHD was clinically assessed and diagnosed according to the DSM-IV-TR criteria,[2] using the Kiddie Schedule for Affective Disorders and Schizophrenia-Present and Lifetime version semi-structured interview template.[18,19] Assessment was performed by experienced psychiatrists specialising in children and adolescents (CAS, ADS, PdC). Parents also completed the ADHD-RS-IV.es questionnaire as part of the diagnostic routine. We excluded patients who had an intelligence quotient of 80 points or less and/or diagnosis of an autism spectrum disorder. We did not exclude patients presenting other comorbidities, such as oppositional defiant disorder or anxiety and/or depressive disorders.

ADHD-RS-IV (parents)
The ADHD-RS-IV[9] is an 18-item scale that evaluates the 18 symptoms listed in the DSM-IV-TR[2] definition of ADHD, which are in turn similar to the current DSM-5 criteria.[1] Each item is scored 0 to 3 on a Likert-type scale (rarely or never; sometimes; often; very often). The total score for the scale is the sum of all item scores. In addition to the total score, 2 subscales reflect the sums of items evaluating inattention and hyperactivity-impulsivity. The Spanish-language version of the scale used in this study (see Table 1) was translated in 1999 by CAS. The translated version was reviewed by an independent translator, and has been updated to take into account the suggestions and difficulties reported by parents, teachers, and other professionals in completing the questionnaire.

The parents of all patients and controls completed the ADHD-RS-IV.es either as part of the diagnostic process (patients) or voluntarily without remuneration (controls). Questionnaires were administered by staff of the child and adolescent psychiatry unit and the paediatrics department at Clínica Universidad de Navarra.

Statistical analysis

We performed a descriptive statistical analysis with measures of central tendency and dispersion (means and standard deviations for quantitative variables, and proportions with 95% confidence intervals [95% CI] for categorical variables). Means were compared using the t test. These analyses were conducted using the SPSS software. We did not differentiate between ADHD subtypes listed in the DSM-IV-TR2 (combined type, predominantly inattentive, and predominantly hyperactive-impulsive). We analysed individual item scores, determining means and standard deviations for the total sample and for each group. We also calculated the Cronbach alpha coefficient and the ordinal alpha coefficient, according to the method described by Zumbo et al.,[20] eliminating one element.

Using the Stata software, a confirmatory factor analysis was performed with structural equation modelling (SEM) in a polychoric correlation matrix (the recommended approach for responses on Likert-type scales with fewer than 7 response options),[20] using the maximum likelihood estimation method. As a basis for comparison, we tested the fit of 2 models: a two-factor model and a hierarchical model, to evaluate the possibility of a two-factor model including a general ADHD factor and specific factors for inattention and hyperactivity-impulsivity, as reported in a previous study.[21] Incremental goodness of fit was measured using the chi-square test, the Tucker—Lewis index (TLI), and the comparative fit index (CFI).[22,23] The root mean square error of approximation (RMSEA) was used as a parsimonious fit measure.[24] Fit is considered to be good in the case of chi-square with degrees of freedom below 2,[25] CFI and TLI equal to or greater than 0.90,[26,27] and RMSEA equal to or smaller than 0.06.[28]

The reliability of the scale was determined using the Stata command *ordalpha*, which is partly based on a modified version of the code by Coveney,[29] to calculate ordinal alpha coefficients for the total scale and subscale scores. These calculations are preferred for Likert-type scales as they are considered to better reflect the degree of concordance.[20] However, to facilitate comparison with the results of other studies in the literature, we also used the SPSS software to calculate Cronbach alpha coefficients[30] for the entire sample for both subscales and for the total scale score.[30] Furthermore, we calculated the average inter-item correlation (r), based on a Pearson correlation matrix and a polychoric correlation matrix, to analyse internal consistency. Average inter-item correlation is recommended for assessing the homogeneity of items, particularly for Likert-type scales with 30 to 40 items. The average inter-item correlation indicates whether items are representative of a single factor; values of 0.15 to 0.50 are recommended.[31]

Finally, we calculated the scale's discriminant validity and predictive value using receiver operating characteristic (ROC) curves. Discriminant validity is an indicator of a scale's capacity to discriminate between the intended subgroups. We calculated 95% CIs for sensitivity and specificity according to the Wilson method,[32] programmed in an SPSS macro. The optimal cut-off point was established with the Youden index (J = sensitivity + specificity − 1), at the point at which the ROC curve is furthest from the diagonal, always using two-tailed P values.

Statistical analysis was performed using SPSS version 15 and Stata version 12.1.

Table 1 The Spanish-language version of the ADHD-RS-IV used in this study.

	Nunca o rara vez	Algunas veces	Con frecuencia	Con mucha frecuencia
IA1 No presta atención a las cosas o comete errores por descuido.				
IA2 Tiene dificultad en mantener la atención en actividades o juegos.				
IA3 No parece escuchar lo que se le está diciendo.				
IA4 Tiene dificultad en completar/terminar tareas o no sigue instrucciones.				
IA5 Tiene dificultad en organizarse (actividades, etc.).				
IA6 Evita actividades que requieren mucho esfuerzo mental (deberes, etc.).				
IA7 Pierde cosas necesarias para actividades o deberes.				
IA8 Se distrae fácilmente por estímulos externos (por el ambiente, etc.).				
IA9 Es olvidadizo de actividades diarias.				
HA1 Mueve manos y pies o se mueve en su asiento.				
HA2 Se levanta en clase o en otras situaciones cuando no debe.				
HA3 Corre o se sube a sitios/cosas cuando se le dice que no lo haga.				
HA4 Tiene dificultad para jugar en silencio.				
HA5 Actúa como si estuviera activado por un motor. «No se le gastan las pilas».				
HA6 Habla excesivamente.				
HA7 Responde abruptamente/impulsivamente antes de terminar la pregunta.				
HA8 Tiene dificultad en esperar su turno en actividades de grupo.				
HA9 Interrumpe a la gente o se mete en conversaciones o juegos.				

Answers are scored 0 to 3 on a Likert-type scale, as follows: ''nunca o rara vez'' (never or sometimes), 0; ''algunas veces'' (sometimes), 1; ''con frecuencia'' (often), 2; ''con mucha frecuencia'' (very often), 3.
HA1-HA9: items corresponding to the 9 symptoms of hyperactivity/impulsivity; IA1-IA9: items corresponding to the 9 symptoms of inattention.

Results

Our sample included 652 children and adolescents: 518 patients (79.73% boys; mean age [SD], 11.40 [3.30] years) diagnosed with ADHD for the first time, and 134 healthy controls (47% boys; mean age, 10.15 [2.97] years).

Target symptoms of ADHD in patients and controls

Patients scored significantly higher than controls for all individual items for inattention and hyperactivity-impulsivity and for the scale as a whole ($P < .001$) (Table 2; Fig. 1).

Discriminant validity of the ADHD-RS-IV.es for differentiating patients from controls

ROC curve analysis revealed a sensitivity of 91.9% and a specificity of 90.3% for discriminating between patients and controls (Fig. 2). The Youden index shows an optimal cut-off point of 17 for the whole sample (17 for boys and 15 for girls; 19 for children and 16 for adolescents) (Fig. 3; Table 2).

Confirmatory factor analysis

The sample was analysed as a whole (patients + controls), independently of DSM-IV-TR[2] subtype, as there is no proven reason in the literature for separating these groups.[28] The two-factor model (inattention, hyperactivity-impulsivity)

showed acceptable fit (CFI = 0.92, TLI = 0.91, RMSEA = 0.07). Confirmatory factor analysis also revealed that all items loaded significantly ($r > 0.30$; $P < .001$) onto the corresponding clinical factors. A strong correlation ($r = 0.62$; 95% CI, 0.57-0.68) was observed between inattention and hyperactivity-impulsivity. Cronbach and ordinal alpha coefficients were high for the total scale score (0.93 and 0.95, respectively) and for both subscales (inattention: 0.93 and 0.95; hyperactivity: 0.91 and 0.94).

Fig. 4 shows the complete SEM diagram.

Discussion

This study evaluates the validity of the ADHD-RS-IV.es and analyses the scale's psychometric properties in a sample of Spanish patients with ADHD and controls. No study published to date has validated a Spanish-language version of the ADHD-RS-IV in such a sample.

Sample size and homogeneity must be taken into account in validation studies for clinical scales. There is no unanimous consensus regarding the minimum number of individuals to be included. The sample size employed in this study meets widely accepted criteria. The total number of participants (652) satisfies the recommendations of Goldberg and Digman[33] that samples should include between 500 and 1000 individuals. The number of participants per questionnaire item (36) meets more demanding criteria.[34] Descriptive analysis of the questionnaire items confirms the

Table 2 Scores for each ADHD-RS-IV.es item in patients (N = 518) and controls (N = 134).

Item	Controls Mean (SD)	Patients with ADHD Mean (SD)	P
Inattention subscale			
Item 1 (IA1): No presta atención a las cosas o comete errores por descuido.	0.63 (0.65)	2.07 (0.77)	<.001
Item 2 (IA2): Tiene dificultad en mantener la atención en actividades o juegos.	0.24 (0.52)	1.52 (0.96)	<.001
Item 3 (IA3): No parece escuchar lo que se le está diciendo.	0.51 (0.59)	1.89 (0.88)	<.001
Item 4 (IA4): Tiene dificultad en completar/terminar tareas o no sigue instrucciones.	0.30 (0.53)	2.12 (0.84)	<.001
Item 5 (IA5): Tiene dificultad en organizarse (actividades, etc.).	0.46 (0.60)	2.17 (0.93)	<.001
Item 6 (IA6): Evita actividades que requieren mucho esfuerzo mental (deberes, etc.).	0.43 (0.62)	2.27 (0.96)	<.001
Item 7 (IA7): Pierde cosas necesarias para actividades o deberes.	0.39 (0.60)	1.60 (1.05)	<.001
Item 8 (IA8): Se distrae fácilmente por estímulos externos (por el ambiente, etc.)	0.73 (0.66)	2.41 (0.71)	<.001
Item 9 (IA9): Es olvidadizo de actividades diarias.	0.45 (0.60)	1.82 (0.94)	<.001
Subscale score for inattention (IA score)	4.13 (3.57)	17.85 (5.48)	<.001
Hyperactivity-impulsivity subscale			
Item 10 (HA1): Mueve manos y pies o se mueve en su asiento.	0.54 (0.75)	1.79 (1.05)	<.001
Item 11 (HA2): Se levanta en clase o en otras situaciones cuando no debe.	0.24 (0.49)	1.41 (1.08)	<.001
Item 12 (HA3): Corre o se sube a sitios/cosas cuando se le dice que no lo haga.	0.37 (0.57)	1.16 (1.1)	<.001
Item 13 (HA4): Tiene dificultad para jugar en silencio.	0.22 (0.50)	0.98 (1.07)	<.001
Item 14 (HA5): Actúa como si estuviera activado por un motor.	0.37 (0.60)	1.26 (1.14)	<.001
Item 15 (HA6): Habla excesivamente.	0.62 (0.86)	1.30 (1.07)	<.001
Item 16 (HA7): Responde abruptamente/impulsivamente antes de terminar la pregunta.	0.42 (0.63)	1.52 (1.05)	<.001
Item 17 (HA8): Tiene dificultad en esperar su turno en actividades de grupo.	0.28 (0.53)	1.23 (1.05)	<.001
Item 18 (HA9): Interrumpe a la gente o se mete en conversaciones o juegos.	0.43 (0.58)	1.48 (1.08)	<.001
Subscale score for hyperactivity-impulsivity (HI score)	3.43 (3.41)	12.11 (7.083)	<.001

ADHD: attention-deficit/hyperactivity disorder; HI: hyperactivity/impulsivity; IA: inattention; SD: standard deviation.
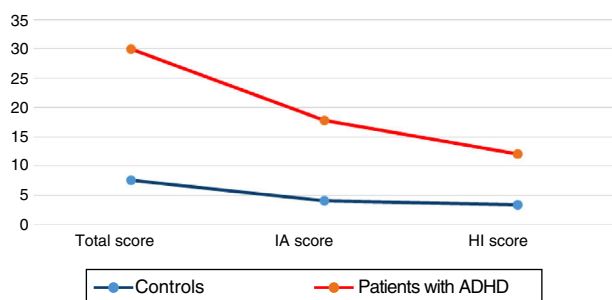

Figure 1 Total and subscale scores for patients and controls on the ADHD-RS-IV.es. ADHD: attention-deficit/hyperactivity disorder; HI: hyperactivity-impulsivity; IA: inattention.

Area under the curve = 0.9694

Figure 2 ROC curve for the ADHD-RS-IV.es questionnaire for the total sample.

tool's ability to differentiate between patients and controls. We detected statistically significant differences (P < .001) between groups for all 18 scale items, total scores, and subscale scores. This demonstrates that the ADHD-RS-IV.es was
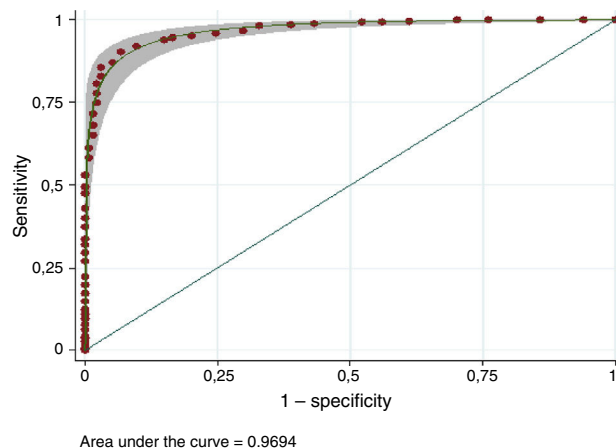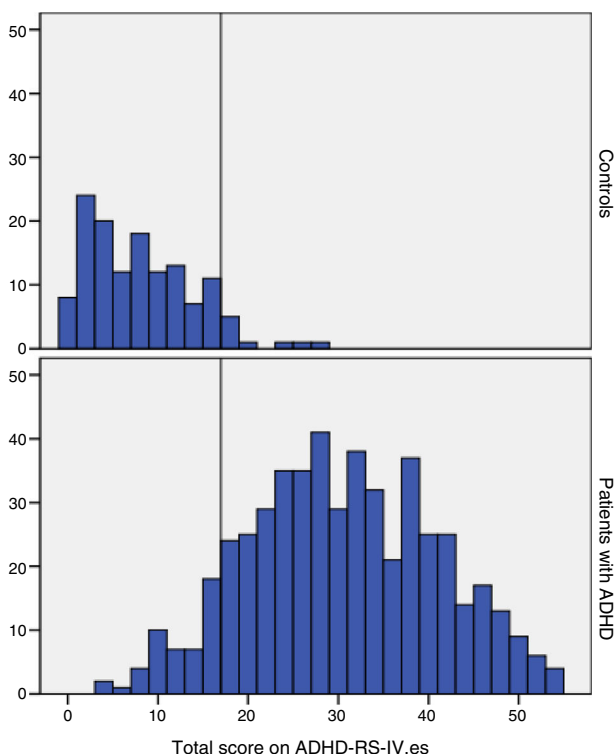
Figure 3 Total ADHD-RS-IV.es scores for patients and controls. The vertical line represents the optimal cut-off point for the sample as a whole (17 points). ADHD: attention-deficit/hyperactivity disorder.

Figure 4 SEM diagram showing a two-factor model and correlations between the 2 subscales (ADHD-RS-IV.es dimensions). Each circle represents one dimension, and each rectangle represents one item (diagnostic criterion). Each circle is joined to the items defining a single dimension and not to the items defining the other dimension (average item-subscale correlation). Absence of an arrow joining an item to a dimension would mean that the factor had a loading of zero for that dimension. Correlations between dimensions and/or items are represented with double-headed arrows (inter-item/inter-subscale correlation coefficient). [a]Average item-subscale correlation (random error). [b]Intersubscale correlation. [c]Inter-item correlation.

able to discriminate between patients with ADHD and controls in our sample. The mean total score for the patient group was 29.96, which is slightly lower than that observed in European (35.8)[16] and multi-national validation studies (41.2).[15] This underscores the importance of obtaining normative data for each country.

An item's discriminating power can be estimated by analysing the corrected item-total correlation coefficient between the item score and the total scale score and between item and subscale scores. Correlation values equal to or greater than 0.25 to 0.30 are considered adequate.[35] As is shown in Fig. 4, all coefficients (factor loadings) are above this range, demonstrating that items are representative of a single factor; this supports the robustness of the scale. The strong correlation ($r > 0.40$) between items or factors implies that both measure very similar aspects. In our model, we found a moderate to strong correlation ($r = 0.62$) between both factors, which demonstrates that both together define a single diagnostic entity. Furthermore, strong positive correlations were observed between various pairs of items: for IA7 (loses things) and IA9 (is forgetful), $r = 0.39$; and for HA8 (has difficulty awaiting turn) and HA9 (interrupts others), $r = 0.41$. These correlations mean that the elimination of one or the other of the items may improve the reliability of the scale. However, in this case, the analysis with elimination of items did not improve goodness of fit; therefore, the model was not modified.
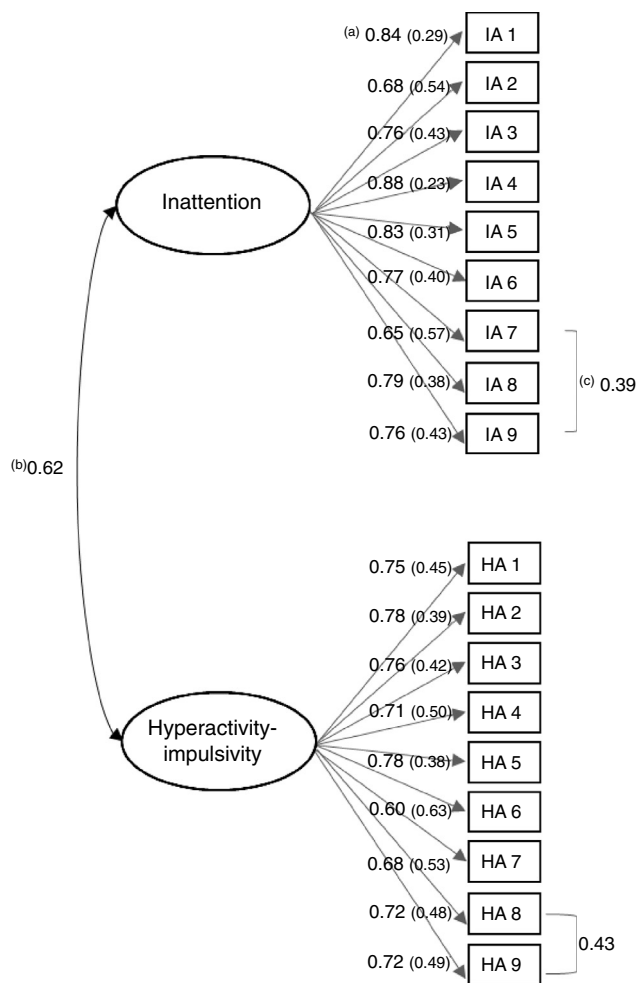
With regard to the factor structure of the ADHD-RS-IV.es, exploratory and confirmatory factor analysis found the two-factor model to be the model best fitting the structure of the scale (CFI = 0.92; TLI = 0.91; RMSEA = 0.07; chi-square = 4.04), scoring over 0.90 for CFI and TLI. This indicates good fit, although some authors recommend values closer to 0.95.[23] RMSEA values below 0.06 indicate good fit and values between 0.06 and 0.08 indicate reasonable fit.[23] In our analysis, chi-square with degrees of freedom was greater than the recommended value of 2.[25] However, this value can be altered in large samples when using the maximum likelihood model; it therefore should

Table 3 Optimal cut-off points (for the total samples and subgroups) and corresponding percentiles from the normative data.

		Cut-off point				Normative data						Cut-off point for the total sample[a] and corresponding percentiles for normative data for the total sample	Cut-off point for subgroups[b] and corresponding percentiles for normative data for each subgroup
		Sensitivity (%)	Specificity (%)	AUC	95% CI	n	Mean (SD)	p90	p93	p95	p98		
Total sample	17[a]	91.9	90.3	0.969	0.96–0.98	134	7.57 (5.83)	15.5	16.55	18	23.90	17 (p93–p95)	
Girls	15[b]	92.4	83.1	0.961	0.94	71	7.42	16	16.96	18	21.68		15(p85–p90)
	17[a]	87.4	88.7		−0.98		(5.67)					17 (p93–p95)	
Boys	17[b]	93.1	92.1	0.969	0.95	63	7.73	15	16.52	17.8	27.44		17(p93–p95)
	17[a]	93.1	92.1		−0.99		(6.04)					17 (p93–p95)	
Children	19[b]	90.8	92.5	0.973	0.96	84	9.25	17	18	19.5	26.6		19(p93–p95)
	17[a]	95.5	86		−0.99		(6.05)					17 (p90–p93)	
Adolescents	16[b]	91.2	95.1	0.986	0.97	50	4.74	11	12.43	13.9	15		16(> p98)
	17[a]	87.8	100		−0.99		(4.14)					17 (>p98)	

95% CI: 95% confidence interval; AUC: area under the receiver operating characteristic curve; SD: standard deviation.

[a] Optimal cut-off point established for the total sample with the highest Youden index value (sensibility + specificity − 1).

[b] Optimal cut-off point established for each subgroups (boys, girls, children, adolescents) with the highest Youden index value (sensibility + specificity − 1).

Table 4 Comparison of our results and those of other studies.

	Internal consistency (Cronbach alpha)		
	Inattention	Hyperactivity-impulsivity	Total
DuPaul et al.[12] (1998)	0.93	0.91	
Zhang et al.[15] (2005)	0.724	0.825	0.795
Döpfner et al.[16] (2006)	0.81	0.87	0.88
Present study	0.937	0.91	0.94

not be used as the only measure of goodness of fit in SEM analysis. Nonetheless, it is beneficial to include it in confirmatory analysis, as it forms the basis of other measures.[33] Despite the high chi-square value for our sample, the excellent values for the other goodness of fit measures (CFI, TLI, and RMSEA) undoubtedly reflect that our scale replicates the factor structure of the original.[9] We also tested the hierarchical model, which showed good fit, contrary to the findings of Ullebø et al.,[21] who observed that the scale fitted a two-factor model with a general ADHD factor.

An instrument's reliability reflects its accuracy in measuring the intended parameter, and is estimated using the Cronbach alpha[30] or the ordinal alpha, which is more accurately for Likert-type scales.[29] Higher alpha values indicate greater internal consistency. We found excellent values for the total scale score (Cronbach alpha, 0.94; ordinal alpha, 0.95) and for the inattention (0.937; 0.95) and hyperactivity-impulsivity subscales (0.91; 0.94); similar values are reported in previous validation studies (Table 3).[14−16]

We also estimated the scale's discriminating power through ROC curve analysis, which is used to calculate sensitivity and specificity for different cut-off points, giving an overall idea of a scale's discriminant capacity. The area under the ROC curve reflects a test's capacity to distinguish between patients and controls: the larger the area under the curve, the greater the discriminating capacity.[36] We established an optimal cut-off point of 17 points for the total scale score; this was associated with high sensitivity and specificity values (92% and 99%, respectively), with an area under the curve of 0.97, indicating good discriminating power. Similar values were observed in subanalyses for age and sex subgroups (Table 4). The cut-off point of 17, obtained for both the total sample and the subsamples, corresponds to percentiles 93 to 95 of the normative data generated for our control group (Table 4). In other words, with this cut-off point, 93% of the total sample would meet criteria for ADHD. Subgroup analysis (boys/girls, children/adolescents) returned similar rates. We also calculated the optimal cut-off points for each subgroup (Table 2), with a higher score for girls than for boys (15 and 19, respectively).

Our study is not without limitations. Firstly, we did not take into account whether the informant was the participant's father or mother, as do some previous studies. Secondly, as the questionnaire was only completed by parents, our data cannot be extrapolated to interpret scale results when completed by teachers. Finally, due to the lack of normative data for the ADHD-RS-IV.es in children and adolescents[17] we included a small control group of 134 individuals. Further studies with larger samples of healthy controls are needed to confirm these preliminary normative data. While our results show that the ADHD-RS-IV.es questionnaire completed by parents can differentiate patients from controls, diagnosis should never be made based exclusively on a questionnaire. It is also essential to interview patients and their parents, to administer questionnaires, to gather information from teachers, and to perform complementary examinations.

Despite these limitations, we believe that the scale presents very good reliability and validity, comparable to those of the original; the same two-factor structure as the original version, proposed in the DSM-IV-TR[2] and the DSM-5 criteria[1]; and clinically acceptable psychometric properties. It is therefore a valid, reliable instrument for evaluating the intensity and frequency of ADHD symptoms.

Conflicts of interest

References

1. American Psychiatric Association (APA). Diagnostic and statistical manual of mental disorders, 5th edition (DSM-5). Washington, DC: American Psychiatric Association; 2013.
2. American Psychiatric Association (APA). Diagnostic and statistical manual of mental disorders, fourth edition, text revised (DSM-IV-TR). Washington, DC: American Psychiatric Association; 2000.
3. DuPaul GJ, Power TJ, Anastopoulos AD, Reid R, McGoey KE, Ikeda MJ. Teacher ratings of attention deficit hyperactivity disorder symptoms: factor structure and normative data. Psychol Assess. 1997;9:436—44.
4. DuPaul GJ, Reid R, Anastopoulos AD, Lambert MC, Watkins MW, Power TJ. Parent and teacher ratings of attention-deficit/hyperactivity disorder symptoms: factor structure and normative data. Psychol Assess. 2016;28:214—25.
5. Anastopoulos AD, Shelton TL. Assessing attention-deficit/hyperactivity disorder. New York: Kluwer Academic/Plenum Press; 2001.
6. Barkley RA, editor. Attention-deficit/hyperactivity disorder: a handbook for diagnosis and treatment. 4th ed. New York: Guilford Press; 2015.
7. DuPaul GJ, Reid R, Anastopoulos AD, Power TJ. Assessing ADHD symptomatic behaviors and functional impairment in school settings: impact of student and teacher characteristics. Sch Psychol Q. 2014;29:409—21.
8. Pelham WE Jr, Fabiano GA, Massetti GM. Evidence-based assessment of attention deficit hyperactivity disorder in children and adolescents. J Clin Child Adolesc Psychol. 2005;34:449—76.
9. DuPaul GJ, Power TJ, Anastopoulos AD, Reid R. ADHD Rating Scale-IV: checklists, norms and clinical interpretation. New York: Guilford Press; 1998.
10. Holland ML, Gimpel GA, Merrell KW. Innovations in assessing ADHD: development, psychometric properties, and factor structure of the ADHD Symptoms Rating Scale (ADHD-SRS). J Psychopathol Behav Assess. 1998;20:307—32.
11. Conners CK. Multi-health systems. 3rd ed. Toronto: Conners; 2008.
12. Swanson J, Nolan W, Pelham WE. The SNAP-IV Rating Scale; 1992 [accessed 14 Feb 2017]. Available in: http://www.adhd.net/snap-iv-instructions.pdf
13. DuPaul GJ, Anastopoulos AD, Power TJ, Reid R, McGoey KE, Ikeda MJ. Parent ratings of ADHD symptoms: factor structure, normative data, and psychometric properties. J Psychopathol Behav Assess. 1998;20:83—102.
14. Power TJ, Doherty BJ, Panichelli-Mindel SM, Karustis JL, Eiraldi RB, Anastopoulos AD, et al. Integrating parent and teacher reports in the diagnostic assessment of ADHD. J Psychopathol Behav Assess. 1998;20:57—81.
15. Zhang S, Faries DE, Vowles M, Michelson D. ADHD Rating Scale IV: psychometric properties from a multinational study as a clinician-administered instrument. Int J Methods Psychiatr Res. 2005;14:186—201.
16. Döpfner M, Steinhausen HC, Coghill D, Dalsgaard S, Poole L, Ralston SJ, et al. Cross-cultural reliability and validity of ADHD assessed by the ADHD Rating Scale in a pan-European study. Eur Child Adolesc Psychiatry. 2006;15 Suppl. 1: S46—55.
17. Servera M, Cardo E. ADHD Rating Scale-IV en una muestra escolar española: datos normativos y consistencia interna para maestros, padres y madres. Rev Neurol. 2007;45: 393—9.
18. Kaufman J, Birmaher B, Brent D, Rao U, Ryan N. Schedule for affective disorders and schizophrenia for school aged children — present and lifetime version (K-SADS-PL). Pittsburgh: University of Pittsburgh Press; 1996.
19. Ambrosini PJ. Historical development and present status of the schedule for affective disorders and schizophrenia for school-age children (KSADS). J Am Acad Child Adolesc Psychiatry. 2000;39:49—58.
20. Zumbo BD, Gadermann AM, Zeisser C. Ordinal versions of coefficients alpha and theta for Likert ratins scales. J Mod App Stat Meth. 2007;6:21—9.
21. Ullebø AK, Breivik K, Gillberg C, Lundervold AJ, Posserud MB. The factor structure of ADHD in a general population of primary school children. J Child Psychol Psychiatry. 2012;53: 927—36.
22. Bentler PM. Comparative fit indexes in structural models. Psychol Bull. 1990;107:238—46.
23. Byrne BM. Structural equation modeling with AMOS: basic concepts, applications and programming. Mahwah: Erlbaum; 2001.
24. Steiger JH, Lind JC. Statistically-based tests for the number of common factors. Paper presented at the annual spring meeting of the Psychometric Society. Iowa City: IA; 1980.
25. Kline RB. Principles and practice of structural equation modeling. New York: Guilford Press; 1998.
26. Boomsma A. Reporting analyses of covariance structures. Struct Equ Model. 2000;7:461—83. https://doi.org/10.1207/S15328007SEM0703_6
27. Browne MW, Cudeck R. Alternative ways of accessing model fit. In: Bollen KA, Long JS, editors. Testing structural equation models. Newbury Park: Sage; 1993. p. 136—62.
28. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model. 1999;6:1—55. https://doi.org/10.1080/107055199095401
29. Coveney J. Calculate alpha after polichoric factor analysis; 2012 [accessed 26 Jun 2017]. Available in: http://www.stata.com/statalist/archive/2012-02/msg00696.html
30. Cronbach LF. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16:297—334.
31. Clark LA, Watson D. Constructing validity: basic issues in objective scale development. In: Kazdin AE, editor. Methodological issues & strategies in clinical research. 3 ed. Washington: APA; 2003. p. 207—31.
32. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat Med. 1998;17:857—72.
33. Goldberg LW, Digman JM. Revealing structure in the data: principles of exploratory factor analysis. In: Strack S, Lorr M, editors. Differentiating normal and abnormal personality. New York: Springer; 1994. p. 216—42.

34. Hair JF, Anderson RE, Tatham RL, Black WC. Multivariate data analysis with readings. 4 ed. Englewood Cliffs: Prentice Hall; 1995.

35. Nunally JC, Bernstein IJ. Teoría psicométrica. Madrid: McGrawHill; 1995.

36. De Irala J, Martínez-González, Sánchez Vilelgas A. Epidemiología clínica. Diagnóstico precoz. In: De Irala-Estévez J, Martínez-González MA, Seguí-Gómez M, editors. Epidemiología aplicada. Barcelona: Ariel; 2008. p. 393—432.