Journal of
Attention Disorders

The Conners Adult ADHD Rating Scales—Short Self-Report and Observer Forms: Psychometric properties of the Catalan version

For Review Only

Journal:	*Journal of Attention Disorders*
Manuscript ID:	JAD-12-02-052.R1
Manuscript Type:	Article
Keywords:	Adult attention deficit hyperactivity disorder, Confirmatory Factor Analysis, Reliability, Validity, Informant

Abstract:	Objective: To validate the Catalan adaptation of the Conners rating scales, short version (Self-Report: CAARS-S:S; and Observer: CAARS-O:S) Method: A community sample of 424 adults responded to the two forms. Confirmatory factor analysis was used to test the dimensional structure. Results: The hypothesized four-factor model (Inattention/Memory Problems, Hyperactivity/Restlessness, Impulsivity/Emotional Lability, and Problems with Self-Concept) presented an adequate fit for both the self-report and observer forms. Reliability was slightly higher for the CAARS-O:S (average α =.78) than for the CAARS-S:S (average α =.75). Test-retest average correlations were: r = .80 (self-reports) and r = .73 (observer ratings). Informant agreement was high at both test (average r = .59) and retest (average r = .61).There were significant gender and age differences. Conclusion: This adaptation of the two short forms of the CAARS-S presents adequate evidence of validity and reliability, and it can therefore be used for diagnostic purposes and cross-cultural comparisons.

SCHOLARONE™
Manuscripts

The Conners Adult ADHD Rating Scales—Short Self-Report and Observer Forms:

Psychometric properties of the Catalan version

Deleted: A Validation Study in a Spanish Community Sample

For Peer Review

The Conners Adult ADHD Rating Scales—Short Self-Report and Observer Forms: Psychometric properties of the Catalan version

Deleted: A Validation Study in a Spanish Community Sample¶

Abstract

Objective: To validate the Catalan adaptation of the Conners rating scales, short version (Self-Report: CAARS-S:S; and Observer: CAARS-O:S) **Method:** A community sample of 424 adults responded to the two forms. Confirmatory factor analysis was used to test the dimensional structure. **Results:** The hypothesized four-factor model (Inattention/Memory Problems, Hyperactivity/Restlessness, Impulsivity/Emotional Lability, and Problems with Self-Concept) presented an adequate fit for both the self-report and observer forms. Reliability was slightly higher for the CAARS-O:S (average α =.78) than for the CAARS-S:S (average α =.75). Test-retest average correlations were: r = .80 (self-reports) and r = .73 (observer ratings). Informant agreement was high at both test (average r = .59) and retest (average r = .61).There were significant gender and age differences. **Conclusion:** This adaptation of the two short forms of the CAARS-S presents adequate evidence of validity and reliability, and it can therefore be used for diagnostic purposes and cross-cultural comparisons.

Deleted: Spanish

Keywords: Attention deficit hyperactivity disorder (ADHD), Conners rating scales, construct validity, reliability, self-report, informant report.

For Peer Review

The Conners Adult ADHD Rating Scales—Short Self-Report and Observer Forms: Psychometric properties of the Catalan version

Deleted: A Validation Study in a Spanish Community Sample ¶

Attention deficit hyperactivity disorder (ADHD) is a neurobehavioural disorder that is usually diagnosed in infancy, childhood or adolescence (DSM-IV-TR, APA, 2000; ICD, WHO, 1995). Three decades ago, ADHD was still believed to be a childhood disorder that disappeared with the onset of adolescence. Since then, however, there has been increasing recognition of the persistence of ADHD across the lifespan (Barkley, Murphy, & Fischer, 2008; Weiss, Trokenberg-Hechtman, & Weiss, 1999; Wender, 1995). In fact, approximately 50% of children with ADHD continue to experience symptoms into adolescence and adulthood. Prevalence rates are between 4% and 5% of the adult population in Europe and the USA (Kessler et al., 2005; Kooij, Buitelaar, van den Oord, Furer, Rijders, & Hodiamont, 2005; Murphy & Barkley, 1996).

Growing scientific evidence suggests that a combination of genetic and environmental factors account for the aetiology of this complex disorder (Singh, 2008). ADHD can cause major disruption not only in the lives of children but also in adults. Despite the apparent similarity between childhood and adult ADHD, the nature of symptoms varies with age (Biederman, Mick, & Faraone, 2000). For example, severe external hyperactivity in childhood may take the form of an internal subjective sense of restlessness in adult ADHD. Impulsivity may manifest as initiating activities or tasks without reading or listening to instructions, or making decisions hastily. Attention problems persist throughout life and manifest as a difficulty in doing things in an order or sequence or as being easily distracted from external and internal stimuli (Barkley et al., 2008; Weiss et al., 1999; Wolf & Wasserstein, 2001).

Diagnosis of ADHD is complex, involving multiple tasks and several sources of information. When diagnosing ADHD in adults it is necessary to take into account the

For Peer Review

frequency and severity of current symptoms, their presence in more than one context (e.g. at home and at work or school) and the prior existence of ADHD symptoms during childhood. Rating scales are useful tools for collecting information on the frequency and severity of ADHD symptoms from various informants who know the person being assessed through different contexts. In recent decades both self-report and observer measures have been developed for the evaluation of adult symptoms of ADHD and the retrospective assessment of childhood ADHD (Adler, Kessler, & Biederman, 2003; Adler, Kessler, & Spencer; 2003; Brown, 1996; Conners, Erhardt, Epstein, Parker, Sitarenios, & Sparrow, 1999; DuPaul, Power, Anastopoulos, & Reid,1998; Kessler et al., 2005; Mehringer, Downey, Schuh, Pomerleau, Snedecor, & Schbiner, 2001; Ward, Wender, & Reimherr, 1993). Although such rating scales cannot be used alone to diagnose ADHD they do provide useful information about the frequency and intensity of ADHD symptoms (Murphy & Adler, 2004).

The Conners Adult ADHD Rating Scales: Psychometric properties

The Conners Adult ADHD Rating Scales (CAARS; Conners, Erhardt, & Sparrow, 1999: Technical Manual) are among the most widely used instruments in the process of diagnosing adult ADHD, and they have been employed in a wide range of clinical settings and in clinical trials. There are three versions of the CAARS, short, long and screening versions, each of which is available in both self-report and observer forms. The short forms of the CAARS (Self-Report form: CAARS-S:S; Observer form: CAARS-O:S) were developed and normed using a sample of non-clinical adults from several sites in the USA and Canada. Dimensionality of the CAARS-S:S and CAARS-O:S was evaluated by confirmatory factor analysis (Conners et al., 1999: Technical Manual). The four-factor structure (Inattention/Memory Problems, Hyperactivity/Restlessness, Impulsivity/Emotional Lability and Problems with Self-Concept) shows a good fit for men and women and across the four

For Peer Review

age groups studied (18-29, 30-39, 40-49 and 50+ years). Cleland Magura, Foote, Rosenblum, and Kosanke (2006) analysed the factor structure of the CAARS-S:S in a sample of 206 substance users, using both exploratory and confirmatory factor analysis. They found that the four-factor structure "was the most parsimonious solution with an adequate degree of fit to the data" (p. 1279), although some items did not load on the expected factor.

The internal consistency of the CAARS-S:S is very good and ranged between .81 (Inattention/Memory Problems, Hyperactivity/Restlessness and Impulsivity/Emotional Lability) and .88 (Problems with Self-Concept) for men, and between .80 (Inattention/Memory Problems and Impulsivity/Emotional Lability) and .85 (Problems with Self-Concept) for women. For the CAARS-O:S, internal consistency ranged between .82 (Hyperactivity/Restlessness) and .89 (Problems with Self-Concept) for men, and between .81 (Hyperactivity/Restlessness) and .88 (Problems with Self-Concept) for women (Conners et al., 1999: Technical Manual).

For the Observer form, test-retest correlations over a two-week interval ranged between .85 (Hyperactivity/Restlessness and Problems with Self-Concept) and .91 (Inattention/Memory Problems; Conners et al., 1999: Technical Manual).

Informant agreement when using the long version of the CAARS Self-Report and Observer forms was studied by Kooij, Boonstra, Swinkels, Bekker, de Noor, and Buitelaar (2008) in a sample of 120 adults with ADHD, who were evaluated by their parents and partners. The results showed a moderate level of agreement between informants, ranging between r = .44 (DSM-IV Inattention Symptoms) and r = .61 (Problems with Self-Concept).

In the standardization sample there were significant gender and age effects for the short versions of the CAARS. On the Self-Report form men scored significantly higher than women on Inattention/Memory Problems and Hyperactivity/Restlessness, whereas women

For Peer Review

scored higher than men on Problems with Self-Concept; there were significant age effects on Impulsivity/Emotional Lability and Problems with Self-Concept. On the Observer form men scored significantly higher than women on Inattention/Memory problems, while there were significant age effects on Hyperactivity/Restlessness and Impulsivity/Emotional Lability (Conners et al., 1999: Technical Manual).

For English-speaking countries there are some well-validated rating scales for assessing ADHD symptoms in adults. However, in our context there are no rating scales that are easy to apply and useful for diagnosing ADHD and for assessing the efficacy of treatment. In light of this the aim of the present study was to test the factor structure and to estimate the reliability (internal consistency and test-retest correlation) and informant agreement of the Catalan version of the CAARS-S: and CAARS-O:S in a large community sample of adults. Catalan is Spain's second most widely spoken language, having more than 13 million speakers in Spain (Catalonia, Valencia and the Balearic Islands), Andorra and some parts of France.

Method

Participants

The sample comprised 424 participants [277 (65.3%) females and 147 (34.7%) males]. The mean age of females was 30.97 years (SD = 13.38, range 18 to 81), while that of males was 33.82 (SD = 12.89, range 18 to 74). For 350 participants the Observer form was completed by someone who knew the participant well (spouse or partner, 53.9%; father/mother, 11.0%; sibling, 11.3%; friend, 19.5%; or child, 4.2%). At the retest stage, data for the Self-Report and Observer forms were available for 116 participants. The sample was recruited in the province of Barcelona through an accidental, non-probability sampling procedure for self-reports, and by means of intentional sampling for observer reports. Both participants and

For Peer Review

observers who responded to the questionnaire were either fluent in Catalan as their first language or understood and spoke Catalan fluently as a second language

Measure

The short versions of the CAARS Self-Report (CAARS-S:S) and Observer forms (CAARS-O:S) (Conners et al, 1999:Technical Manual) are paper-and-pencil measures designed to assess ADHD symptoms in adults. Both forms comprise a 26-item rating scale scored on a 4-point Likert scale (0 = 'not at all or never', 1 = 'just a little, once in a while', 2 = 'pretty much, often', and 3 = 'very much, very frequently'). The wording and order of items on the CAARS-S:S and CAARS-S:O is the same and only differs in the style of questions. Items on the CAARS-S:S are formulated in the first person singular (e.g. *I interrupt others when talking*), whereas items on the CAARS-S:O are formulated in the third person singular (e.g. *The person being described interrupts others when talking*)

The CAARS-S takes approximately ten minutes to complete and includes four factor-derived subscales (Inattention/Memory Problems, 5 items; Hyperactivity/Restlessness, 5 items; Impulsivity/Emotional Lability, 5 items, and Problems with Self-Concept, 5 items) and two created scales (the ADHD Index, 12 items that best distinguish ADHD adults from non-clinical adults; and the ADHD Inconsistency Index, 8 pairs of items that have similar content and which can be used to identify random or careless responding).

Adaptation

The adaptation of the short versions of the CAARS was carried out by the Department of Personality, Assessment and Psychological Treatment of the University of Barcelona and the Psychiatry Department of the Vall d'Hebron University Hospital. Multi-Health Systems

For Peer Review

Inc. (MHS) gave permission for this adaptation and confirmed that it conformed to their normative standards.

The translation and adaptation procedure was as follows (Brislin, 1986; Brislin, Lonner & Thorndike, 1973; Candell & Hulin, 1987): (1) direct translation of the original English scales into Catalan by two expert translators with knowledge of psychological assessment and psychopathology; (2) comparison of these two translations to assess any differences of interpretation; (3) back translation from Catalan into English, by two different expert translators; (4) comparison of the direct and back translations by bilingual individuals in order to check the conceptual and semantic equivalence of the item statements; and (5) drawing up the definitive version that is presented in this study. The back translation was sent to MHS for review. After making minor corrections, MHS accepted the translated version.

Data analysis

The dimensional structure of the short versions of the CAARS was tested by confirmatory factor analysis (CFA), using the twenty items that were retained in the CFA by Conners et al. (1999). The first step involved testing the four-factor structure (Inattention/Memory Problems, Hyperactivity/Restlessness, Impulsivity/Emotional Lability and Problems with Self-Concept) proposed by the scales' developers. Competing models were also analysed and compared with the abovementioned four-factor model. Specifically, a comparison was made with: 1) a one-factor, second-order model on which the four primary factors loaded, which would account for the global construct ADHD, the root of the four specific factors; and 2) a one-dimensional, first-order model on which all the items loaded. The three model specifications are shown in Figure 1. In all the factor structures tested, each item was specified to load on one factor only, and measurement errors were not allowed to correlate.

For Peer Review

Deleted: ¶
INSERT FIGURE 1 ABOUT HERE¶

The analysis was performed using the LISREL 8.8 program (Jöreskog & Sörbom, 2006a). Given the ordinal nature of the data the maximum likelihood robust method was used for estimation, while polychoric correlations and their corresponding asymptotic covariance matrices were previously generated by means of the PRELIS 2.8 program (Jöreskog & Sörbom, 2006b). The goodness-of-fit of the models shows us how well each theoretical model tested matches the data. This degree of fit was assessed with the following indices: the Satorra–Bentler scaled chi-square (S-B χ^2), the root mean square error of approximation (RMSEA) and its relative confidence interval, the non-normed fit index (NNFI), the comparative fit index (CFI) and the Akaike information criterion (AIC). Indicators of a good fit are that S-B χ^2 is not significant, that NNFI and CFI have values above .95, and that the RMSEA value does not exceed .05 (Kaplan, 2000). The AIC index has a comparative purpose and the model with the lower value shows the best fit.

Reliability was measured in terms of internal consistency, test-retest correlations and informant agreement. Internal consistency was evaluated by means of Cronbach's α. Following Haertel (2006), Cronbach's α values were considered as follows: $.60 \leq r < .80$ adequate; $.80 \leq r < .85$ good; and $r \geq .85$ excellent. Test-retest and informant agreement were assessed by means of Pearson's correlations. Correlation values were considered as follows: $< .3$ = small; .3 to .5 = moderate; and $\geq .5$ = large (Cohen, 1988).

In order to analyse age and gender differences the scores for each factor were derived from the sum of the scores for the items that loaded on each factor (Russell, 2002). Differences in factor ratings were calculated for age groups and gender by means of ANOVA. All the calculations were performed using SPSS version 19.

For Peer Review

Results

Dimensional structure

Table 1 shows the fit indexes achieved by the three models tested, for both self- and observer-report data. The results of the S-B $\chi 2$ statistic showed that the fit was poor for all the models (p <.01). However, given that this index is affected by sample size, by the degrees of freedom in the model and by discrepancies in the normality of the data, authors such as Kaplan (2000) and MacCallum and Austin (2000) recommend that the assessment of fit be based on the alternative criteria that were listed in the data analysis section. These indices show that the one-factor, first-order model should be discarded due to a lack of fit: specifically, it yielded values of χ^2/df > 6, RMSEA > .10, and NNFI and CFI indexes < .95, as well as an AIC that was clearly higher than that of the other models. The other two specified models (the four-factor model and the second-order model) present good fit values for all the indices considered: χ^2/df < 2, NNFI and CFI indexes > .95, and RMSEA values close to .05 (in all cases the .05 value falls within the 90% confidence interval). This means that these two models are plausible for both the Self-Report (CAARS-S:S) and Observer (CAARS-O:S) forms. However, comparison of the AIC indices shows that the first-order (four-factor) model offers the best fit for both types of data (self-report and observer), and it therefore represents a better approximation to reality. All the loadings of this model were statistically significant (p<.01), which means that all the indicators are relevant for defining the corresponding domain.

INSERT TABLE 1 APPROXIMATELY HERE

For Peer Review

Reliability

Table 2 shows the internal consistency (Cronbach's α coefficients) and test-retest correlations (after a one-month interval) for both self and observer ratings and the informant agreement at test and retest in relation to the four factors obtained in the CFA. Internal consistency is generally adequate for observer ratings (mean α = .780) and for self-reports (mean α = .750). The highest internal consistency coefficient for both forms (Self and Observer ratings) corresponds to the factor Problems with Self-Concept (.835 and .862 for Self and Observer ratings, respectively), while the lowest coefficient corresponds to the factor Hyperactivity/Restlessness (.710 and .651 for self and observer ratings, respectively). We also used the W statistic (Feldt, 1969) to test the differences between the alpha coefficients in both samples. The results confirmed the hypothesis that the two alphas were not different at the 1% level, given that the critical value of F (with 423 and 349 df) was 1.27.

Test-retest correlations are high for both self-reports (mean r = .799, p <.001) and observer ratings (mean r = .734, p < .001). The lowest test-retest correlations correspond to the subscale Problems with Self-Concept (for both the Self-Report and Observer forms).

Informant agreement is high at both the test stage (mean r = .593, p<.001) and retest (mean r = .612, p<.001). The lowest agreement at the test stage corresponds to the factors Inattention/Memory Problems and Problems with Self-Concept, while at retest the least agreement is found for Problems with Self-Concept. Application of Fisher's Z test to the correlation coefficients of the Self and Observer versions that were obtained in the analysis of test-retest and informant agreement revealed no significant differences in any of the comparisons.

INSERT TABLE 2 APPROXIMATELY HERE

For Peer Review

Age and gender differences

Tables 3 and 4 show the means and standard deviations for the CAARS-S:S and CAARS-O:S, separately for gender and age groups. A series of (gender x age group) ANOVAs was conducted with each of the CAARS factors as the dependent variable (for both the Self-Report and Observer forms). These analyses revealed significant gender differences on the factor Problems with Self-Concept: Self-Report form, $F(1,342) = 7.258$, $p = 0.007$, $\eta^2 = 0.018$; and Observer form, $F(1,416) = 7.513$, $p = 0.006$, $\eta^2 = 0.021$. Females had higher mean raw scores than males on both the Self-Report and Observer forms. On the Self-Report form, age had a significant effect on the factors Inattention/Memory Problems [$F(3,416) = 4.775$, $p = .003$, $\eta^2 = .033$] and Hyperactivity/Restlessness [$F(3,413) = 14.283$, $p < .001$, $\eta^2 = .094$]. For Inattention/Memory Problems participants aged 18-29 years scored significantly higher than did those aged 40-49 years, there being no significant differences between the other age groups. For Hyperactivity/Restlessness there were significant differences between the scores of participants aged 18-29 years and those of participants in both the 40-49 and 50+ age groups, with the youngest group scoring higher.

On the Observer form there were significant age differences for Inattention/Memory Problems [$F(3,341) = 4.390$, $p = .005$, $\eta^2 = .037$], Hyperactivity/Restlessness [$F(3,341) = 4.158$, $p < .007$, $\eta^2 = .035$] and Impulsivity/Emotional Lability [$F(3,340) = 2.707$, $p = .045$, $\eta^2 = .023$]. For Inattention/Memory Problems and Hyperactivity/Restlessness participants aged 18-29 scored significantly higher than did those aged 40-49. For Impulsivity/Emotional Lability participants aged 50+ (especially males) scored significantly higher than did those aged 40-49 years.

Clinically significant ADHD symptoms are indicated by T-scores ≥70. Based on self-report ratings, 8 men (5.4%) and 11 women (3.97%) from the sample had scores ≥70 on the four

For Peer Review

factor scales (Inattention/Memory Problems, Hyperactivity/Restlessness, Impulsivity/Emotional Lability and Problems with Self-Concept).

INSERT TABLES 3 AND 4 APPROXIMATELY HERE

Discussion

Rating scales are highly useful instruments in the process of diagnosing ADHD and for evaluating the efficacy of treatments. In the USA and Europe the Conners scales are probably the most widely used instrument in the clinical and research contexts. Of course, when any measurement instrument is used in a context other than that in which it was developed it is necessary to check that its validity and reliability are adequate and comparable across cultures. In this regard the aim of the present study was to analyse the psychometric properties of the short Catalan versions of the CAARS Self-Report (CAARS-S:S) and Observer (CAARS-O:S) forms in a large community sample of adults.

Deleted: validate

Deleted: the dimensional structure of

The results of the CFA show that it would not be adequate to derive a total score for the CAARS-S:S and CAARS-O:S, and they corroborate the dimensional validity of the four factors. The optimum model, among the three tested, is that proposed by the scales' developers, i.e. the one containing four primary factors: Inattention/Memory Problems, Hyperactivity/Restlessness, Impulsivity/Emotional Lability and Problems with Self-Concept. Given that these scales are widely used in our context the confirmation of this model is important in that it offers a guarantee that scores are being adequately interpreted; it also justifies the use of this dimensional structure for evaluating ADHD symptoms and associated behaviours in our population.

Deleted: discriminant

Deleted: among

Internal consistency is high, although the values are slightly lower than those obtained in the standardization sample (Conners et al., 1999: Technical Manual). These lower values may

For Peer Review

be due to the fact that the present sample was smaller than the one used for standardization, which could lead to less variability. Test-retest correlations, with a one-month interval, are significant and, in most cases, are moderate or high; they are also higher for the Self-Report than the Observer form. The highest test-retest correlations were obtained for the Inattention/Memory Problems subscale, while the lowest corresponded to the Problems with Self-Concept subscale (for both Self-Report and Observer forms). The values obtained in the present study are similar to those reported for the standardization sample (Conners et al., 1999: Technical Manual), although it should be borne in mind that the original standardization was based solely on the long version of the Self-Report form (CAARS-S:L), whereas here we used the short versions of both the Self-Report and Observer forms. Nonetheless the present data support the reliability of the scores obtained by participants and informants in evaluating the symptoms of ADHD, and the use of these scores in the process of clinical diagnosis.

The assessment of ADHD in adults is a complex process since there is no diagnostic test as such. Therefore, in order to increase the accuracy and precision of the diagnosis it is necessary to gather information from a range of instruments (interviews, rating scales, psychometric and laboratory tests and observational measures) and from different informants. Rating scales, completed by someone who knows the person being assessed, can provide highly useful information as regards improving diagnostic accuracy (Pelham, Fabiano, & Masseti, 2005), and the degree of expected agreement between different sources is important for the diagnosis. Research with self-referred clinical samples generally indicates that informants (spouses, partners, parents or siblings) report more symptoms than do adults diagnosed with ADHD (Barkley, Fischer, Smallish, & Fletcher, 2002; Katz, Petscher, & Welles, 2009; Kooij,et al., 2008), although the study by Zucker, Morris, Ingram, Morris and Bakerman (2002), involving 281 college students with academic difficulties and

For Peer Review

who self-referred for assessment, found that while informants reported more significant inattentive symptoms there were no differences in terms of hyperactive/impulsive symptoms. In community samples, adults reported more ADHD symptoms than did other informants (Murphy & Schachar, 2000). Several studies have found strong correlations between participant and informant ratings of current ADHD symptoms when using a DSM-IV-based checklist in clinical (r = .39 to .42; Kooij et al., 2008; r = .69; Murphy & Barkley, 1996; r = .55 to .57; Zucker et al., 2002), community (r = .59 to.70; Murphy & Schachar, 2000) and clinical and control samples (r = .67; Barkley, Knouse, & Murphy, 2011). In the present study, self and observer ratings were positively correlated. Informant agreement ranged between moderate and high, and was higher for the Hyperactivity/Restlessness and Inattention/Memory Problems subscales than for the Problems with Self-Concept subscale. The values obtained were of similar magnitude to those reported by Kooij et al. (2008) for long versions of the CAARS, and by Barkley et al. (2011), Murphy & Barkley (1996), Murphy & Schachar (2000) and Zucker et al. (2002) for scales based on DSM-IV ADHD symptoms.

Between 4% and 5% of our sample showed significant symptoms of ADHD that would be consistent with the presence of the disorder. These percentages are similar to those found in other studies with community samples of adults (Barkley et al., 2008).

As in the standardization sample, gender differences were observed in relation to the Problems with Self-Concept subscale. Women scored higher than men across all age groups on both the Self-Report and Observer forms.

Age was also a differential factor as regards the scores on different subscales, with scores generally being lower as age increased. The youngest age group (18-29 years) scored highest, and scores then decreased across the subsequent age groups (30-39, 40-49 and 50+) on all but one factor, namely Impulsivity/Emotional Lability. On this factor, scores

For Peer Review

decreased from the 18-29 years group up to participants aged 40-49, but then rose slightly in the 50+ group, an effect that was mainly due to the subgroup of men. Behaviours associated with impulsivity, irritability and sudden mood changes were more intense among the oldest age group. This greater frequency and intensity of impulsive behaviours was particularly evident on the Observer form. As this phenomenon was not observed in relation to the standardization sample it may be due to particular characteristics of the present sample, this being an aspect that should be investigated in future research.

In summary, this study is the first to analyse construct validity, reliability, and age and gender differences for the short Self-Report and Observer forms of the CAARS in a large community sample of Catalan adults. The findings show that the Catalan adaptation of the test has adequate validity. Furthermore, the results confirm the original scale structure, which means that this adaptation is adequate for use in the process of diagnosing ADHD and for monitoring the efficacy of treatment. This is important because these scales are used internationally and it is necessary to have instruments whose factor structure is comparable cross-culturally so as to facilitate research and enable results to be compared.

Deleted: Spanish

For Peer Review

References

Adler, L., Kessler, R, & Biederman, J. (2003). *Adult ADHD Investigator Symptom Rating Scale*. New York: New York University School of Medicine.

Adler, L., Kessler, R., & Spencer, T. (2003). *Adult Self-Report; Scale. ASR-V1.1. Screener*. New York: World Health Organization.

American Psychiatric Association, APA (2000) *Diagnostic and statistical manual of mental disorders: DSM-IV-TR* (text revision) Washington, DC: Author.

Barkley, R. A., Fischer, M., Smallish, L., & Fletcher, K. (2002). The persistence of attention-deficit/hyperactivity disorder into young adulthood as a function of

reporting source and definition of disorder. *Journal of Abnormal Psychology, 111*, 279–289.

Barkley, R. A., Knouse, L. E., & Murphy K. R. (2011).Correspondence and Disparity in the Self- and Other Ratings of Current and Childhood ADHD Symptoms and Impairment in Adults With ADHD. *Psychological Assessment, 23*, 437-446.

Barkley, R. A., Murphy, K. R., & Fischer, M. (2008). *ADHD in adults. What the science says*. New York: Guilford Press.

Biederman, J., Mick, E., & Faraone, S. V. (2000). Age-dependent decline of symptoms of attention deficit hyperactivity disorder: impact of remission definition and symptom type. *American Journal of Psychiatry, 157*, 816-8.

Brislin, R. W. (1986). The wording and translation of research instruments. In W. J. Lonner & J. W. Berry (Eds.), *Field methods in cross-cultural psychology.* (pp. 137-164) Newbury Park, CA: Sage.

Brislin, R. W., Loner, W. J., & Thorndike, R. M. (1973). *Cross-cultural research methods.* New York: Wiley

Brown, T. E. (1996). *Brown Attention-Deficit Disorden Scales*. San Antonio, TEX: Psychological Corporation.

Candell, G. L., & Hulin, C. L. (1987). Cross-language and cross-cultural comparisons in scale translations. *Journal of Cross-Cultural Psychology, 17,* 417-440.

Cleland, C., Magura, S., Foote, J., Rosenblum, A., & Kosanke, N. (2006). Factor structure of the Conners Adult ADHD Rating Scale (CAARS) for substance users. *Addictive Behavior, 31,* 1277-1282.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition). New Jersey: Lawrence Erlbaum.

For Peer Review

Conners, C. K., Erhardt, D., Epstein, J. N., Parker, J. D. A., Sitarenios, G., & Sparrow, E. (1999). Self-ratings of ADHD symptoms in adults I: Factor structure and normative data. *Journal of Attention Disorders, 3*,141-151.

Conners, C. K., Erhardt, D. & Sparrow, E. (1999). CAARS. *Adult ADHD Rating Scales. Technical Manual*. Toronto: Multi-Health System.

DuPaul, G. J., Power, T.J., Anastopoulos, A.D., & Reid R. (1998). *ADHD Rating Scale-IV: Cheklist, Norms, and Clinical Interpretation.* . New York, Guilford Press.

Feldt, L. S. (1969). A test of the hypothesis that Cronbach's alpha or Kuder-Richardson coefficient twenty is the same for two tests. *Psychometrika, 34*, 363-373.

Haertel, E. H. (2006). Reliability. In R. L. Brennan (Ed.), *Educational measurement* (pp. 65–110). Wesport, CT: American Council on Education and Praeger Publishers.

Jöreskog, K., & Sörbom, D. (2006a). *LISREL 8.8 for Windows* [Computer software]. Lincolnwood, IL: Scientific Software International, Inc.

Jöreskog, K., & Sörbom, D. (2006b). *PRELIS 2.8 for Windows* [Computer software]. Lincolnwood, IL: Scientific Software International, Inc.

Kaplan, D. (2000). *Structural equation modelling: Foundations and extensions.* Thousand Oaks, CA: Sage.

Katz, N., Petscher, Y., & Welles, T. (2009). Diagnosing Attention-Deficit Hyperactivity Disorder in College Students: An Investigation of the Impact of Informant Ratings on Diagnosis and Subjective Impairment *Journal of Attention Disorders, 13*, 277-283.

Kessler, R. C., Adler, L., Ames, M., Demler, O., Faraone, S., Hiripi, E., Howes, M. J., Jin, R., Secnik, K., Spencer, T., Ustun, T. B., & Walters E .E. (2005). The World Health Organization Adult ADHD Self-Report Scale (ASRS): a short screening scale for use in the general population. *Psychological Medicine,* 35, 245-256

For Peer Review

Kooij, J. J. S., Buitelaar, J. K., van den Oord, E. J., Furer, J. W., Rijders, C. A. T., & Hodiamont, P. P. G. (2005). Internal and external validity of attention-deficit/hyperactivity disorder in a population based sample of adults. *Psychological Medicine, 15*, 817-827.

Kooij, J. J. S., Boonstra, A. M., Swinkels, S. H. N., Bekker, E. M., de Noor, I., & Buitelaar, J. K. (2008). Reliability, validity, and utility of instruments for self-report and informant-report concerning symptoms of ADHD in adults patients. *Journal of Attention Disorders, 11*, 445-458.

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modelling in psychological research. *Annual Review of Psychology*, *51*, 201-226.

Mehringer, A. M., Downey, K. K., Schuh, L. M., Pomerleau, C. S., Snedecor, S. M., & Schbiner, H. (2001). The Assessment of Hyperactivity and Attention: Development and preliminary validation of a brief self-assessment of adult ADHD. *Journal of Attention Disorders*, *5,* 223-31.

Murphy, K. R. & Adler, L. A. (2004). Assessing attention-deficit /hyperactivity disorder in adults: Focus on rating scales. *Journal of Clinical Psychiatry, 65* (suppl 3), 12-17.

Murphy, K. & Barkley, R. A. (1996). Prevalence of DSM-IV symptoms of ADHD in adult licensed drivers: Implications for clinical diagnosis. *Journal of Attention Disorders, 1*, 147-161.

Murphy, P. & Schachar, R. (2000). Use of self-ratings in the assessment of symptoms of attention deficit hyperactivity disorder in adults. *American Journal of Psychiatry, 157,* 1156-1159.

Pelham, W. E., Fabiano, G. A., & Massetti, G. M. (2005). Evidence-based assessment of Attention Deficit Hyperactivity Disorder in children and adolescents. *Journal of Clinical Child Psychology, 34*, 449-476.

For Peer Review

Russell, D.W. (2002). In search of underlying dimensions: the use (and abuse) of factor analysis in Personality and Social Psychology Bulletin. *Personality and Social Psychology Bulletin*, *28*(12), 1629-1646. Singh, I. (2008). Beyond polemics: science and ethics of ADHD. *Nature Reviews, Neuroscience, 9, 957-964.*

Ward, M. Ff., Wender, P. H., & Reimherr, F. W. (1993). The Wender Utah Rating Scale: An Aid in the Retrospective Diagnosis of Childhood Attention Deficit Hyperactivity Disorder. *American Journal of Psychiatry, 150*, 885-890.

Weiss, M., Trokenberg-Hechtman, L., & Weiss, G. (1999). *ADHD in adulthood. A guide to current theory, diagnosis, and treatment*. Baltimore: The Johns Hopkins University Press.

Wender PH: *Attention-Deficit Hyperactivity Disorder in adults*. New York, Oxford University Press, 1995.

Wolf, L., y Wasserstein, J. (2001). Adult ADHD: Concluding thoughts. *Annals of the New York Academy of Sciences, 931*, 396– 408.

World Health Organization. (1995). *The ICD-10 classification of mental and behavioural disorders*. Geneva: World Health Organization.

Zucker, M., Morris, M. K., Ingram, S. M., Morris, R. D., & Bakerman, R. (2002). Concordance of self-and informant ratings of adults' current and childhood attention-deficit/hyperactivity disorder symptoms. *Psychological Assessment, 14*, 379-389.

For Peer Review

Table 1. CFA goodness-of-fit statistics

Model	χ^2	$\chi^{2/}$df	RMSEA	90% CI	NNFI	CFI	AIC
SELF-REPORT FORM							
Four first-order factors	320.34	1.941	0.047	0.040-0.055	0.98	0.98	412.34
Four factors loading on a higher-order factor	327.09	1.970	0.048	0.040-0.056	0.98	0.98	415.09
One factor	1053.55	6.197	0.11	0.10-0.12	0.90	0.91	1135.54
OBSERVER FORM							
Four first-order factors	304.29	1.855	0.049	0.041-0.058	0.98	0.98	396.29
Four factors loading on a higher-order factor	325.58	1.961	0.052	0.044-0.061	0.98	0.98	413.58
One factor	1380.99	8.123	0.14	0.14-0.15	0.84	0.85	1460.99

CFA= Confirmatory Factor Analysis; χ^2 = Satorra-Bentler scaled chi-square, df = degrees of freedom; RMSEA = root mean square error of approximation; CI = confidence interval; NNFI = non-normed fit index; CFI = comparative fit index; AIC = Akaike information criterion

For Peer Review

Table 2. Internal consistency (Cronbach's α), test-retest correlations and informant agreement for self and observer ratings, and their corresponding differences.

	Internal consistency			Test-retest			Informant agreement		
Factors	Self (n= 424)	Observer (n = 350)	Feldt' W	Self (n = 116)	Observer (n = 116)	Fisher' Z	Test (n = 350)	Retest (n = 116)	Fisher' Z
Inattention/Memory Problems	.740	.770	.885	.831	.761	1.44	.567	.629	-.89
Hyperactivity/Restlessness	.710	.651	1.203	.805	.762	.84	.609	.688	-1.26
Impulsivity/Emotional Lability	.714	.838	.566	.795	.752	.80	.607	.644	-.56
Problems with Self-Concept	.835	.862	.836	.765	.663	1.57	.588	.489	1.29

Note. All correlation are significant at p<.001.

For Peer Review

Table 3. Means and standard deviations (in brackets) for the four factors of the CAARS-S:S

	18-29 years		30-39 years		40-49 years		50+ years	
Factors	Males	Females	Males	Females	Males	Females	Males	Females
Inattention/Memory Problems	4.22 (3.09)	3.75 (3.12)	4.23 (3.36)	3.42 (2.37)	2.14 (2.23)	3.00 (2.71)	3.75 (3.42)	2.73 (2.98)
Hyperactivity / Restlessness	4.75 (3.21)	3.96 (2.72)	3.12 (2.69)	2.69 (2.54)	2.28 (2.34)	2.25 (1.96)	2.00 (2.55)	3.23 (3.24)
Impulsivity / Emotional Lability	3.18 (3.79)	3.95 (2.95)	3.08 (3.26)	3.54 (2.90)	2.69 (2.00)	2.95 (2.21)	4.20 (3.46)	3.04 (3.04)
Problems with Self-Concept	3.60 (3.25)	5.14 (3.18)	3.58 (3.35)	3.85 (3.66)	2.59 (2.41)	4.36 (3.02)	3.40 (3.60)	4.04 (2.87)
N	72	168	26	26	29	56	20	27

For Peer Review

Table 4. Means and standard deviations (in brackets) for the four factors of the CAARS-O:S

	18-29 years		30-39 years		40-49 years		50+ years	
Factors	Males	Females	Males	Females	Males	Females	Males	Females
Inattention/Memory Problems	3.57 (2.91)	3.16 (2.66)	3.80 (3.59)	2.65 (3.07)	1.66 (1.91)	2.27 (2.61)	3.79 (3.67)	2.15 (2.71)
Hyperactivity / Restlessness	3.61 (3.47)	3.40 (2.52)	2.81 (2.87)	3.00 (2.74)	2.38 (2.04)	2.00 (2.08)	3.40 (3.08)	2.73 (2.25)
Impulsivity / Emotional Lability	3.43 (3.62)	4.11 (3.05)	2.42 (2.50)	4.35 (3.17)	3.28 (2.43)	2.56 (2.34)	5.65 (4.33)	3.31 (3.18)
Problems with Self-Concept	2.44 (2.86)	4.01 (3.32)	2.92 (3.07)	3.65 (3.47)	1.72 (1.91)	3.35 (3.28)	3.05 (2.74)	3.30 (3.48)
N	61	107	25	26	29	55	19	27

For Peer Review