

# Factorial validity, reliability of assessments and prevalence of ADHD behavioural symptoms in day and residential treatment centres for children with behavioural problems

E.M. SCHOLTE, I.A. VAN BERCKELAER-ONNES, Department of Educational Sciences, Leiden University, The Netherlands

J.D. VAN DER PLOEG, Dutch Institute for Psychological Research (Nippo), Amsterdam, The Netherlands

**ABSTRACT** *This study uses the attention deficit/hyperactivity disorder (ADHD) symptom ratings of professional care workers to estimate the prevalence of ADHD symptoms among children in day treatment centres (N = 162) and residential treatment centres (N = 195) in Holland. Although further research is needed, the study supports the suggestion that such ratings can add to reliable diagnostic outcomes when assessing the behavioural symptoms of ADHD in children in the centres. It is estimated that nearly a fifth of the children in such centres exhibit the symptoms of ADHD in the judgement of professional care workers.*

*Model testing using confirmatory factor analysis favours a dimensional behavioural model that comprises all the three constitutional symptom dimensions of ADHD (inattention, hyperactivity and impulsivity) instead of the two-factor model as used in the DSM-IV (inattention and hyperactivity/impulsivity). However, the differences of fit between both models were only small and the hyperactivity and impulsivity factors were highly correlated. This suggests that, in practice, a two-factor model can also be appropriate. The issue of whether a two-factor or a three-factor model is more appropriate thus remains unsolved in this study.*

**Key words:** ADHD, symptoms, residential care centres

## Introduction

Attention deficit/hyperactivity disorder (ADHD) is one of the most common disorders in childhood. It comprises inattentive, impulsive and hyperactive behavioural symptoms in children. Prevalence studies suggest that all over the world 3% to 9% of all young children in the general population exhibit behavioural symptoms related to this disorder (Swanson et al., 1998). Children with ADHD usually display a range of other problems as well. They often have learning difficulties, without having a classified learning disorder, and are known for their aggressive and oppositional behaviour (Barkley, 1998; Munden and Arcelus, 1999; Wenar and Kerig, 1999). The prognosis of the long-term adaptive emotional and behavioural development of these children is poor (Swanson et al., 1998). Some ADHD children outgrow or overcome their problems with the assistance of their parents and teachers. However, often the

problems are unmanageable without professional aid and the children are referred to healthcare institutions for the treatment of children with behavioural difficulties. Here successful treatment always starts with a correct diagnosis of the disorder.

Diagnosing ADHD is not an easy task, however. Although there is a growing consensus that some kind of neurological and/or biochemical process must be involved, the causes of ADHD remain elusive (Grodzinsky and Barkley, 1999; Barkley, 1999; Hobwenko, 1999). Until confirmed medical markers are found, the diagnosis can only be established at the behavioural level. To assess the symptoms of ADHD correctly at the behavioural level some conditions must be met. In the first place a valid set of symptoms is needed, as differences in labels and definitions result in different diagnostic criteria. Fortunately, after decades of different operational definitions, the DSM

and ICD, in their most recent versions, reached an almost identical set of 18 symptoms. However, differences in decision rules still remained (Swanson et al., 1998).

In the second place, a robust identification of the core syndromes is required. Many exploratory factor analytic studies have identified two behavioural subdimensions of ADHD (an inattention factor and a hyperactive/impulsive factor) (see for example: Achenbach and Edelbrock, 1978; Bauermeister et al., 1995; Hinshaw, 1987; Lahey and Carlson, 1991; Loeber, Lahey and Thomas, 1991; Brito, Pinto and Lins, 1995; Healy et al., 1993; Pelham, Evans, Gnagy, and Greenslade, 1992). However, recent studies have also suggested that approaches using the three dimensions of inattention, hyperactivity and impulsivity separately cover the structure of ADHD better (Pillow et al., 1998; Gomez et al., 1999) and thus question the two-factor structural model.

In the third place, assessments of the behavioural symptoms must be reliable. In the past the diagnostic criteria were criticized because diagnosticians often differed in their assessment of the same children (Quay, 1979; Taylor, 1988; Nelson-Gray, 1991; Swanson et al., 1998). To redress this, rating scales were developed to assess the hyperkinetic syndrome in children (Achenbach, 1991; Conners, 1970; Barkley, 1998). However, these scales do not include the ADHD symptoms as listed in the most recent versions of the ICD and the DSM. As such these older instruments cannot be used to assess ADHD according to the most widely accepted criteria (Pelham et al., 1992). Recent studies suggested that DSM-IV-based rating scales can be used to specify validly and reliably the main and subdimensional symptoms of ADHD (Baumgartel et al., 1995; Pillow et al., 1998; Gomez et al., 1999).

In Holland the most difficult children with ADHD are usually placed in day treatment centres (DTC) or in residential treatment centres (RTC). Here specifically educated professional care workers take care of the children and aim to provide an optimal basic therapeutic climate for these children (see for example Scholte and Van der Ploeg, 2000).

The studies mentioned above using DSM-IV-oriented rating scales were based on the ratings of parents and teachers. They do not report findings for professional care workers. As these professionals work with children with behavioural difficulties on a professional and daily

basis, it is conceivable that they will rate the behavioural symptoms of children in a different way from teachers and parents. It is, however, very important to know whether professional care workers are able to rate the symptoms of ADHD reliably and validly. These professionals must not only assess the ADHD symptoms correctly at the beginning of the treatment to choose a potential successful therapeutic climate for the children but also during the treatment itself. A comparison of robustly made ADHD ratings at the beginning and in follow-ups will give a reliable and valid insight into the behavioural progress of the ADHD children during the treatment. Any progress found during the treatment can be interpreted as supporting the view that the conditions provided were successful and can be continued whereas a lack of progress suggests that the treatment was not successful and should be replaced by a more suitable therapeutic regime.

This study used professional care workers' ratings of a scale comprising the 18 symptoms of ADHD as specified by the DSM on a sample of children admitted to residential care. The study had three aims. The first was to examine the factorial validity of the behavioural ratings of ADHD by the professional care workers in professional care centres. The second was to assess the reliability of the assessments made by the professional care workers. The third was to estimate prevalence rates of the behavioural symptoms of ADHD according to the judgements of the professional care workers in a sample of (semi)residential children as for the DSM-IV core syndromes (for example, the symptom clusters of the combined type, the inattentive type and the hyperactive/impulsive type).

The practical aim of the study was to take a first step in the development of a valid, reliable and easy-to-use instrument for professional care workers to assess the behavioural symptoms of the ADHD-disorder according to the most recent and widely used behavioural criteria for ADHD.

## **Method**

### *Participants*

In the Netherlands the professional treatment of ADHD children can take place in day treatment centres (DTC) or in residential treatment centres (RTC). Children in DTCs stay at their family homes and receive professional treatment during working hours.

Most of these centres serve only younger children, up to the age of 12. Children in RTCs are placed out of home in a residential institution and receive professional care and treatment on a 24-hour basis. It is estimated that, of the total Dutch youth population of 4 million children, about 4,000 are placed in DTCs yearly and about 2,000 children are placed in RTCs (Scholte, 1997).

Two groups of children were available for this study. The first group was a random sample of 162 children who were admitted to DTCs. The sample was drawn by taking the last four newly admitted children of 39 DTCs randomly selected from all 70 DTCs in the Netherlands. Due to workloads some centres delivered only three cases, whereas others delivered five. The DTC sample consisted of 73% boys and 27% girls, the mean age of the children was 6.7 years (standard deviation 3.4, range 4–16 years).

The second sample was a random sample of 195 children admitted to RTCs. This sample was drawn by taking the last 20 newly admitted children from the total of 10 nationwide working RTCs in the Netherlands. This sample consisted of 69% boys and 31% girls. The mean age of the children was 11.8 years (standard deviation 2.4, range 6–18 years). Both samples reflected the populations of the care areas rather well (Scholte, 1997).

The total sample of this study thus consisted of 357 children. It comprised 29% girls and 71% boys. The age range of the children was four to 18 years. The mean age was 9.4 years, the standard deviation 4.0 years and the range 4–18 years.

### Measures

*Attention-Deficit/Hyperactivity Disorder (ADHD) rating scale* The behavioural symptoms of ADHD were measured using a rating scale comprising all 18 symptoms listed in the DSM-IV. The respondents were instructed to rate each item on a five-point scale as follows: 0 – not at all; 1 – just a little (incidentally); 2 – pretty much (at least monthly); 3 – much (at least weekly); 4 – very much (at least daily). The appendix lists the items used in the rating scale.

### Analysis

*Factorial validity* In the DSM tradition the core symptoms of ADHD have always been composed of

inattention, impulsivity and hyperactivity. In DSM-III (American Psychiatric Association, 1980) the three core symptoms were declared separate dimensions. DSM-III-R (American Psychiatric Association, 1987) provided a single symptom list indicating one dimension, whereas DSM-IV (American Psychiatric Association, 1994) provides two subsyndromes (inattention and hyperactivity/impulsivity) and one main syndrome (the combined type of inattention and hyperactivity/impulsivity). In this study the factorial validity was investigated in terms of which of these three models most accurately captures the ratings of the professional raters. The three models tested were the one-factor model as specified by the DSM-III-R, the two-factor model as specified by the DSM-IV and the three-factor model as specified by the DSM-III.

The models were investigated by using a confirmatory instead of an exploratory factor analysis (Long, 1988; Bentler, 1992; Byrne, 1994). In exploratory factor analysis the correlations determine what factors will be identified. This approach discards any theoretical considerations of factor composition beforehand. In a confirmatory factor analysis an alternative approach is used. First a set of concurrent factor models is specified based on theoretical considerations. Next, the fit of each model is computed and the models are compared as for the best fit (Anderson and Gerbing, 1988).

### Reliability

To assess the behavioural symptoms of ADHD correctly, the symptom ratings made by the professional care workers must be minimally affected by error. It would, for example, be unacceptable if one professional should conclude that a certain child has the behavioural symptoms of the disorder if another would reach the opposite conclusion when assessing the same child.

*Reliability* refers to the extent to which a test, or any measurement procedure, yields consistent results in repeated trials. Three different methods are usually minimally applied to assess the reliability of a test or rating scale (Nelson-Gray, 1991; Carmines and Zeller, 1992): the internal consistency method, the test-retest method and the inter-rater reliability method.

The term *internal consistency* refers to the extent that the items of a test or rating scale relate to each other and the internal consistency method yields an estimate of these interrelationships – Cronbach alpha, a coefficient between zero and one.

The *test-retest method* yields an estimate of this consistency over time. Reliable assessment implies that administration of the test or rating procedure twice by the same person should yield comparable results. The test-retest estimate is usually expressed as the intra-class correlations between both measures (Deyo et al., 1992). The method is problematic, however. For example, the symptoms to be assessed can change over time because the child was treated. Moreover, the ratings of the second assessment can easily be influenced by the ratings of the first assessment because raters can remember the ratings they made on the first occasion. It is usually recommended to choose a time compromise that reduces both these effects. A test-retest period of about two to three weeks between both test administrations seems optimal. This is not long enough for a substantial treatment effect, but it is long enough to rule out memory influences.

The *inter-rater reliability* refers to the extent that two independent raters reach the same conclusion about the same child. This is usually accomplished by two raters administering the test or rating scale independently of each other for the same child at the same point in time. Here the intra-class correlation between test or rating scale administrations is a suitable measure to express the consistency of judgement between raters (Deyo et al., 1992).

#### *Prevalence*

In this study the term *prevalence* refers to the rate the behavioural symptoms of the ADHD disorder can be witnessed in the group of children studied according to the judgements of the professional care workers. In this study this prevalence rate was estimated by calculating the percentages of children in DTCs and in RTCs expressing the three core behavioural syndromes of ADHD: the percentage of children showing the behavioural symptoms of the inattentive type and the hyperactive/impulsive type respectively, and the combined type at a clinical level stipulated by the behavioural criteria of the DSM-IV.

#### *Procedure*

The DSM-IV-based ADHD rating scale was filled in by 60 care professionals working in the DTCs and in the RTCs. These professionals were in charge of daily care for the children. The ratings were made at the end of the first month after the admittance of the children to

the treatment centre. To explore the test-retest reliability of the rating scales, the respondents were asked to fill in second rating scales for the same children two weeks afterwards. To explore the inter-rater reliability of the rating scales, two respondents were asked to administer the rating scales for the same children independently at the same time of the day. The data gathering in the RTCs took place in 1997 and in the DTCs in 1998.

## **Results**

#### *Factorial validity*

The first model to be tested comprised all 18 DSM-IV symptoms together. This model reflects the factorial structure of the ADHD-disorder as formulated in the DSM-III-R and the DSM-IV combined type. In the second model the two factors in accordance with the DSM-IV were modelled. In the third model the three factors of inattention, hyperactivity and impulsivity were modelled using the symptom lists for these subdimensions as stipulated in the DSM-III.

The variance-covariance structure of all three models was subjected to a confirmatory factor analysis using robust maximum likelihood estimation in EQS (Bentler, 1992). The results of this analysis are presented in Table 1.

A widely used method to determine the model fit is the  $\chi^2$  test. In general it is assumed that significant  $\chi^2$  values represent poor fits.

As the entries in Table 1 for the total sample show, none of the three models fit the data satisfactorily according to this criterion. The value of the  $\chi^2$  'goodness of fit' test is, however, strongly determined by the number of cases in the sample, with large numbers of cases inflating the  $\chi^2$ . In this case Bentler (1988) and also Byrne (1994) recommend using indices that are less dependent on the sample size: the comparative fit-index or CFI (Bentler, 1988) and the robust variant of the CFI that corrects for deviations from multivariate normality (RCFI) (Bentler and Dijkstra, 1989). Both the CFI and the RCFI range from zero to one. Models with a (R)CFI of 0.95 and above are usually considered to represent the observed covariance matrix satisfactorily (Bentler, 1992). In addition to these fit indices, the root mean square error of approximation can be calculated. The RMSEA reflects the lack of fit of a model. Smaller values thus represent a better fit. Models with

**Table 1.** Fit indices of the ADHD-factor models composed of one, two and three factors

Factor model:	Model with 1 factor	Model with 1. Inattention 2. Hyperactivity/impulsivity	Model with 1. Inattention 2. Hyperactivity 3. Impulsivity
Total sample (N=357):			
$\chi^2$ (Df/P-value)	659.5 (135/<.01)	603.7 (132/<.01)	534.1 (131/<0.01)
RCFI	0.92	0.95	0.95
RMSEA	0.10	0.10	0.09
AIC	389.4	339.7	272.1
4-12 year sample (N=207):			
$\chi^2$ (Df/P-value)	446.9 (135/<.01)	430.8 (132/< 0.01)	405.1 (131/< 0.01)
RCFI	0.92	0.93	0.96
RMSEA	0.11	0.10	0.10
AIC	176.9	166.8	143.0
13-18 year sample (N=150):			
$\chi^2$ (Df/P-value)	467.4 (135/< 0.01)	330.7 (132/< 0.01)	317.2 (131/< 0.01)
RCFI	0.89	0.95	0.97
RMSEA	0.13	0.10	0.09
AIC	197.4	66.7	55.2
DTC- sample (N=195):			
$\chi^2$ (Df/P-value)	435.8 (135/< 0.01)	429.0 (132/< 0.01)	392.1 (131/< 0.01)
RCFI	0.90	0.92	0.92
RMSEA	0.12	0.12	0.11
AIC	165.8	165.0	130.1
RTC- sample (N=162):			
$\chi^2$ (Df/P-value)	509.6 (135/< 0.01)	439.1 (132/< 0.01)	415.9 (131/< 0.01)
RCFI	0.91	0.94	0.94
RMSEA	0.12	0.11	0.09
AIC	239.6	175.1	153.9

$\chi^2$ =Chi-squared value; Df=Degrees of freedom; P-value=significance level; RCFI=Robust comparative fit index; RMSEA=root mean square error of approximation; AIC=Akaike's information criterion.

values of 0.08 or smaller are usually considered to represent the data well (Browne and Cudeck, 1993; McCallum et al., 1996). In addition to this Akaike's information criterion (AIC) was calculated. This criterion takes into account both the statistical goodness of fit and the number of parameters that have to be estimated to achieve that degree of fit. The model that produces the minimum value may be considered potentially the most useful (Dunn, Everitt and Pickles, 1993).

Table 1 shows that the one-factor model fitted the data the least well. In fact this model just fails to meet

the minimum requirements of model fit as set out above. The two-factor model consisting of the inattention symptoms as the first factor and the hyperactivity/impulsivity symptoms as the second factor gives a better representation of the data. The model composed of all three dimensions seems to represent an even better fit. It produces the smallest RMSEA and AIC, suggesting that this model is potentially the most useful.

To test the robustness of these findings for the younger and the older children, as well as across both professional care settings, the above calculations were

also done within the age groups of the 4–12-year-old children and the 13–18-year-old children, and also within the groups of DTC and RTC children. In all groups the one-factor model fitted the data the least well, the two-factor model fitted the data second best and the three-factor model showed a slightly better fit than the two-factor model.

It must be noted, however, that the difference between the models with two and three factors is disputable, as no huge differences in the explained RCFI-covariance between both models was found in the various tests.

Table 2 shows the correlation coefficients between the factors in the two-factor model and in the three-factor model. These coefficients were computed to examine how the different symptom dimensions are related to each other. As can be noted, in the two-factor model both dimensions were highly correlated. In the three-factor model all three dimensions were also highly correlated, but the hyperactivity and impulsivity dimensions correlated higher than both these dimensions correlated with the inattention dimension, in particular in the DTC sample and in the sample with younger children. Thus there was a tendency for hyperactivity and impulsivity to be more strongly associated overall than inattention and hyperactivity, or inattention and impulsivity.

#### Reliability

Table 3 presents the reliabilities for the rating scales based on the DSM-IV symptom list ordered according to the factors of the three ADHD-models tested here.

The first three columns of the table present the internal consistencies for the combined DTC and RTC sample, the DTC sample and the RTC sample respectively.

Table 3 shows that the internal consistencies of the factor based ADHD-rating scales for all models are 0.82 or above. Widely used diagnostic scales usually require alphas of 0.80 or above (Carmines and Zeller, 1993). Thus for each model, the internal consistencies of the ADHD-symptoms rating scales based on the factors of the respective models are satisfactorily high.

Due to workload considerations the test-retest and also the inter-rater reliability methods were only obtained for smaller subsamples randomly drawn from the RTC-sample. The test-retest period was 16.7 days (standard deviation 6.7 days). The intra-class correlations between the subsequent measures turned out to be satisfactorily high for all the rating scales based on the respective factors of all three models. They are all 0.89 or above, suggesting good test-retest characteristics for all scales, as intra-class correlations of 0.75 are described as ‘excellent’ (Landis and Koch, 1977; Fleiss, 1981). The inter-rater reliabilities were also satisfactorily high. They were all 0.82 or above.

The findings for the internal consistencies, the test-retest and the inter-rater reliabilities were also tested for the children under 12 years versus those above 12 years and for males versus females. In all these subgroups they turned out to be satisfactorily high (all above 0.80).

These findings suggest that ordering of the DSM-IV symptoms list of ADHD into rating scales leads to measurements sufficiently reliable for the assessment of the respective main or subdimensions of the behavioural

**Table 2.** Correlations between the factors in the two- and three-factor models

	Total sample (N=357)		4-12 sample (N=207)		13(18 sample (N=150)	
Two-factor model	Inattention		Inattention		Inattention	
Hyperactivity/impulsivity	0.80		0.77		0.84	
Three-factor model	Inattention	Impulsivity	Inattention	Impulsivity	Inattention	Impulsivity
Inattention	–	0.71	–	0.65	–	0.77
Hyperactivity	0.68	0.85	0.63	0.85	0.74	0.82
	DTC sample (N = 162)		RTC sample (N = 195)			
Two-factor model	Inattention		Inattention			
Hyperactivity/impulsivity	0.83		0.78			
Three-factor model	Inattention	Impulsivity	Inattention	Impulsivity		
Inattention	–	0.63	–	0.80		
Hyperactivity	0.62	0.83	0.76	0.85		

**Table 3.** Reliabilities of DSM-IV based ratings using ADHD models with one, two and three factors

	(N=357) <sup>3</sup>	Internalconsistency <sup>1</sup> (N=162) <sup>4</sup>	(N=195) <sup>5</sup>	Test-retest <sup>2</sup> (N=190) <sup>5</sup>	Inter-rater <sup>2</sup> (N=109) <sup>5</sup>
One-factor model:					
ADHD combined	0.94	0.93	0.96	0.96	0.95
Two-factor model:					
Inattention	0.88	0.85	0.93	0.91	0.94
Hyperactivity/impulsivity	0.92	0.91	0.94	0.93	0.93
Three-factor model:					
Inattention	0.82	0.80	0.88	0.89	0.91
Hyperactivity	0.92	0.90	0.93	0.94	0.95
Impulsivity	0.88	0.87	0.90	0.91	0.92

<sup>1</sup> Cronbach's Alpha; <sup>2</sup> Intra-class correlations; <sup>3</sup> Total sample; <sup>4</sup> Day-treatment sample; <sup>5</sup> Residential-treatment sample

symptoms, irrespective of whether one uses a model with one (the combined AD/HD type), two (attention deficit and hyperactivity/impulsivity) or three (attention deficit, hyperactivity and impulsivity) dimensions.

#### Prevalence

The prevalence rates of the different types of ADHD symptoms were computed for the two-dimensional classification of the DSM-IV. In line with the DSM-IV criteria, the inattentive type consisted of children with at least six symptoms from the inattention list of the DSM-IV, whereas the hyperactive-impulsive type consisted of children with at least six symptoms from the hyperactive-impulsive list of the DSM-IV. The combined type consisted of those children showing at least six symptoms of both the inattentive and the hyperactive/impulsive symptoms list. As in other studies (Baumgartel et al., 1995; Pelham et al., 1992; Gomez et al., 1999) ratings of 'much' and 'very much' were interpreted as indicating the presence of the symptom.

The prevalence rates are presented in Table 4. As the prevalence rates for boys and girls in general populations differ considerably (see, for example, Swanson et al., 1998) also in this study the prevalence rates were analysed gender. Table 4 further shows the 95% confidence intervals for the prevalence rates in the various (sub)samples.

For the total sample (of both the residential and day treatment children) the prevalence rate of the behavioural symptoms for all ADHD types was 21%. The rates for the combined types, the inattentive and the hyperactive/impulsive types were 10%, 7% and 4% respectively. Calculating relative risks for males and

females using a 95% confidence interval showed that males in care run 1.2 to 3.5 times more risk of having the symptoms of ADHD than females in care.

For the DTC the prevalence rate for all ADHD types together was 20. Here the rates for the combined types, the inattentive and the hyperactive/impulsive types were 8%, 5% and 7% respectively. Calculating the relative male-female risks showed here that males in DTCs run between one and four times more risk of having the symptoms of ADHD than females in DTC.

In the RTCs the prevalence rate for all ADHD types together was 24%. The rates for the combined types, the inattentive and the hyperactive/impulsive types were 9%, 5% and 9% respectively. Calculating the relative risks showed here that males in RTC run 0.9 to five times more risk of having the symptoms of ADHD than females in RTC.

#### Conclusion and discussion

Recent studies have suggested that DSM-IV-based rating scales filled in by parents and teachers can be used to assess the behavioural symptoms of ADHD validly and reliably (Baumgartel et al., 1995; Pillow et al., 1998; Gomez et al., 1999). However, no known studies have found out yet whether such ratings can also be validly and reliably made by professional care workers in day care and residential treatment centres. Moreover, no known studies have estimated the prevalence of the behavioural symptoms of ADHD in such centres using DSM-IV-based rating scales filled in by professional care workers. To answer these questions, professional care takers working with behaviourally difficult children in day care and residential

**Table 4.** Estimated prevalence rates of different types of DSM-IV defined ADHD symptoms based on ratings of professional care workers in day care and residential treatment centres in Holland

Total sample	All (N=357)		Boys (N=252)		Girls (N=105)	
	N	%	N	%	N	%
Inattention	25	7 ( $\pm 2.5$ ) <sup>a</sup>	22	9 ( $\pm 3.5$ )	3	3 ( $\pm 3.1$ )
Hyperactivity/impulsivity	16	4 ( $\pm 2.0$ )	12	5 ( $\pm 2.6$ )	4	4 ( $\pm 3.7$ )
Combined type	35	10 ( $\pm 3.1$ )	29	11 ( $\pm 3.8$ )	6	5 ( $\pm 4.1$ )
Total (all types)	76	21 ( $\pm 4.3$ )	63	25 ( $\pm 5.3$ )	13	12 ( $\pm 6.2$ )

  

DTC sample	All (N=162)		Boys (N=118)		Girls (N=44)	
	N	%	N	%	N	%
Inattention	12	7 ( $\pm 3.9$ )	10	9 ( $\pm 5.1$ )	2	5 ( $\pm 6.4$ )
Hyperactivity/impulsivity	8	5 ( $\pm 3.3$ )	7	6 ( $\pm 4.2$ )	1	2 ( $\pm 4.1$ )
Combined type	13	8 ( $\pm 4.1$ )	11	9 ( $\pm 5.1$ )	2	5 ( $\pm 6.4$ )
Total (all types)	33	20 ( $\pm 6.1$ )	28	24 ( $\pm 7.1$ )	5	12 ( $\pm 9.6$ )

  

RTC sample	All (N=195)		Boys (N=144)		Girls (N=49)	
	N	%	N	%	N	%
Inattention	15	9 ( $\pm 4.0$ )	12	9 ( $\pm 4.6$ )	1	2 ( $\pm 3.5$ )
Hyperactivity/impulsivity	8	5 ( $\pm 3.0$ )	5	4 ( $\pm 3.0$ )	3	5 ( $\pm 5.4$ )
Combined type	15	9 ( $\pm 4.2$ )	18	13 ( $\pm 5.4$ )	4	6 ( $\pm 5.8$ )
Total (all types)	46	24 ( $\pm 5.9$ )	35	26 ( $\pm 7.1$ )	8	13 ( $\pm 8.4$ )

<sup>a</sup> 95% confidence interval

treatment centres in Holland were asked to fill in DSM-IV based rating scales.

Three concurrent models were investigated for reliability and validity: a model with one rating scale comprising all the 18 DSM-IV symptoms, a model comprising two rating scales with the symptoms arranged as understood in the DSM-IV (inattention and hyperactivity/impulsivity) and a model comprising three rating scales with the symptoms arranged as understood in the DSM-III (inattention, hyperactivity and impulsivity).

For the total sample, as well as for the children under 12 years of age and the children above 12 years of age, and for the daycare and residential treatment samples, the two- and the three-factor models indicated a significantly better fit of the ratings compared to the one-factor model, whereas the three-factor model indicated a slightly better fit than the two-factor model. This matches the findings of recent studies comparing the factorial structure of the hyperactive disorder by means of comparative factor analysis of the DSM-IV ratings by parents and teachers (Pillow et al., 1998; Gomez et al., 1999). These studies also reported

that an ordering of the DSM-IV behavioural symptoms according to three factors (inattention, hyperactivity and impulsivity) produced better fits than ordering according to two factors (inattention and hyperactivity/impulsivity combined), with the differences in fit being only minor.

As for the reliability of the scales it was found that the internal consistencies, the test-retest reliabilities and the inter-rater reliabilities of all scales in all three models amply met the standards set for rating scales widely used for diagnostic purposes. All reliabilities were above 0.80, a finding that can be evaluated as excellent (Carmines and Zeller, 1992).

This suggests that the ADHD symptom ratings made by professional care workers in day and residential treatment centres can add to a reliable assessment of the behavioural symptoms of ADHD. It must be noted, however, that assessing the behavioural symptoms of ADHD this way must not be equated with reaching the diagnosis of ADHD as additional decision rules must be applied for this.

The finding that the scales did equally well no matter whether the ordering of the symptoms followed the



one-factor approach (the combined type), the two-factor approach (attention deficit and hyperactivity/impulsivity) or the three-factor approach (attention deficit, hyperactivity and impulsivity) further suggests that it is foremost a matter of factorial validity (and not reliability) that determines which behavioural model represents the ADHD-disorder the best. It must be noted that the differences in fit between the three-factor model and the two-factor models were relatively minor. In the three-factor model the correlations between the hyperactive scale and impulsivity scale were strong and higher than the correlations between the inattentive scale and the hyperactive and impulsivity scales in both the day care and residential treatment samples. According to others (Gomez et al., 1999) such a finding would suggest that the two-factor model represents the most appropriate model for the organization of the diagnostic symptoms of ADHD. In addition to this reasoning for parsimony would suggest adopting the simplest of all valid models, thus also opting for the two-factor model. However, in the context of competitive model testing reasoning for parsimony is not a valid argument (Chow, 1996). From a statistical point of view the best fitting model is the one to be preferred, which would argue in favour of adopting the three-factor model. Moreover, Akaike's information criterion reaches the lowest value for the three-factor model, suggesting that the three-factor model is the most useful. Usually contrary arguments can be decided by allowing for the results of other studies. However, many other studies also favour both the two- and the three-factor models. For example, studies using exploratory factor analysis are largely in favour of the two-factor approach (Bauermeister et al., 1992; Lahey et al., 1988; Pelham et al., 1992; Brito et al., 1995; Holland et al., 1998), whereas studies using comparative factor analysis are slightly in favour of the three-factor approach (Gomez et al., 1999), in particular when corrections are made for co-occurring conduct disorders (CD) and oppositional defiant behaviour (OD) (Pillow et al., 1998). Such contrary findings mean that the issue of the best organization of the behavioural symptoms of ADHD is rather unresolved.

On the basis of the ratings of the professional care workers the prevalence rate of all types of ADHD symptoms for both gender groups together was 20% for the DTCs and 24% for the RTCs. These overall prevalence rates are higher than the overall rates of 5% to

10% found in general youth populations (Swanson et al., 1998; Kalverboer, 1996). Within a 95% confidence interval, two to five times more boys than girls showed the symptoms of ADHD. This compares with the ratios of 2:1 till 6:1 usually found for ADHD in youth populations (Weiss, 1996). However, caution must be exercised in interpreting the prevalence rates. First, as has been noted, the DSM-IV diagnosis of ADHD requires additional criteria, like the presence of the symptoms before the age of seven. These additional criteria were not included while assigning children into the various diagnostic groups. The prevalence rates thus must not be interpreted as the prevalence of the diagnosis ADHD. Second, due to the small sample size, in particular, the estimates of the behavioural symptoms of ADHD for the girls in the day and residential treatment centres, one must allow for the rather wide confidence interval.

In concluding, this study supports the suggestion that ADHD-symptoms ratings made by professional care workers can add to reliable diagnostic outcomes as for the assessment of the behavioural symptoms of ADHD by children in day and residential treatment centres. Using this assessment method nearby a fifth of the children in such centres express the symptoms of ADHD. However, some limitations of the study must be mentioned.

In the first place, whether the behavioural symptoms of ADHD are organized according to the two dimensions of the DSM-IV (inattention and hyperactivity/impulsivity), or whether a model with the three dimensions of inattention, hyperactivity and impulsivity is more appropriate, remains unresolved in this study. To address this a more thorough theoretical clarification is probably needed of what makes up the nature of the hyperactive/impulsive part of ADHD (Nelson-Gray, 1991).

In the second place the predictive validity of the assessments made by the professional care workers was not investigated in this study. Although the factorial validity findings suggest that the judgements of the professional care workers reflect the two- and three-factor DSM models about equally well, and the reliability findings for the scales are satisfactorily high, the study has to be completed by showing that the assessments made by the professional care workers can predict the diagnosis of ADHD satisfactorily. This can be done by comparing the judgements of the professional care workers with a clinical diagnosis of the

same children independently made by a child psychiatrist or child psychologist. In addition it must be shown that the judgements of the professional care workers compare with the judgements of parents and teachers, as these are the judgements the clinician can normally obtain.

Besides these points, the high co-occurrence of ADHD and other disorders, like conduct disorders (CD), oppositional defiant disorder (OD) and mood-and-anxiety disorders repeatedly found in studies (Angold et al, 1999) raises the question whether ADHD exists in practice as a pure disorder. Be this as may, at the individual clinical level such comorbidities must first satisfactorily be assessed and ruled out before the diagnosis 'ADHD' can correctly be made (American Psychiatric Association, 1994). We address these issues in a further study.

## References

- Achenbach TM, Edelbrock CS. The classification of child psychopathology: a review and analysis of empirical efforts. *Psychological Bulletin* 1978; 85: 1275–301.
- Achenbach, TM. Manual for the Child Behaviour Checklist/4-18. Burlington: University of Vermont, Department of Psychiatry, 1991.
- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 3 edn. Washington DC: APA, 1987.
- American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 4 edn. Washington DC: APA, 1994.
- Anderson JC, Gerbing DW. Structural equation modeling in practice: a review and recommended two-steps approach. *Psychological Bulletin* 1988; 103: 411–23.
- Angold A, Costello EJ, Erkalani A. Comorbidity. *Journal of Child Psychology and Child Psychiatry* 1999; 40: 57–87.
- Barkley RA. Attention-Deficit Hyperactivity Disorders. A Handbook for Diagnosis and Treatment. New York: Guilford, 1988.
- Barkley RA. Theories of attention-deficit/hyperactivity disorder. In HC Quay, AE Hogan (eds) *Handbook of Disruptive Behaviour Disorders*. New York: Kluwer Academic/Plenum Publishers, 1999.
- Bauermeister JJ, Alegria M, Bird H, Rubio-Stipic M, Canino G. Are attentional-hyperactivity deficits unidimensional or multidimensional syndromes? Empirical findings from a community survey. *Journal of the American Academy of Child and Adolescent Psychiatry* 1992; 32: 423–32.
- Baumgartel A, Wolraich ML, Dietrich M. Comparison of diagnostic criteria for attention deficit disorders in a German elementary school sample. *Journal of the American Academy of Child and Adolescent Psychiatry* 1995; 34: 629–38.
- Bentler PM. Comparative fit indexes in structural models. *Psychological Bulletin* 1988; 107: 238–46.
- Bentler PM, Dijkstra T (1989) Efficient estimation via linearization in structural models. In PR Krishnaiah (eds) *Multivariate Analysis VI*. Amsterdam: North-Holland Publishing, 1989.
- Bentler PM. On the fit of models to covariances and methodology to the bulletin. *Psychological Bulletin* 1992; 112: 400–4.
- Bentler PM. EQS Structural Equations Program Manual. Encino CA: Multivariate Software, Inc, 1995.
- Brito GN, Pinto RCA, Lins MFC. A behavioural assessment scale for attention deficit disorder in Brazilian children based on DSM-III-R criteria. *Journal of Abnormal Child Psychology* 1995; 23: 509–20.
- Browne MW, Cudeck R (1993) Alternative ways of assessing model fit. In KA Bollen, JS Long (eds) *Testing Structural Equation Models*. Newbury Park CA: Sage, 1993.
- Burns GL, Walsh JA, Patterson DR, Holte CS, Sommers-Flanagan R, Parker CM. Internal validity of the disruptive behaviour disorder symptoms: implications from parents ratings for a dimensional approach to symptom validity. *Journal of Abnormal Child Psychology* 1997; 25: 307–19.
- Byrne BM. *Structural Equation Modeling with EQS and EQS/Windows*. London: Sage, 1994.
- Carmines EG, Zeller RA. Reliability and Validity Assessment. In M Lewis-Beck (ed.) *Basic Measurement*. London: Sage, 1994.
- Chow SL. *Statistical Significance: Rationale, Validity and Utility*. London: Sage, 1996.
- Conners KC. Symptom patterns in hyperkinetic, neurotic and normal children. *Child Development* 1970; 41: 667–82.
- Deyo RA, Diehr P, Patrick DL. Reproducibility and Responsiveness of Health Status Measures. *Controlled Clinical Trials*, 1991; 12: 142s–58s.
- Dunn G, Everitt B, Pickles A. *Modelling Covariances and Latent Variables using EQS*. London: Chapman-Hall, 1993.
- Fleiss JL. *Statistical Methods for Rates and Proportions*. New York: Wiley, 1981.
- Gomez R, Harvey J, Quick C, Scharer I, Harris G. DSM-IV AD/HD: Confirmatory factors models, prevalence, and gender and age differences based on parent and teacher ratings of Australian primary school children. *Journal of Child Psychology and Psychiatry*, 1999; 40: 265–74.
- Grodzinsky GM, Barkley RA. Predictive power of frontal lobe tests in the diagnosis of attention-deficit hyperactivity disorder. *Clinical-Neuropsychologist* 1999; 13: 12–21.
- Healy JM, Newcorn JH, Halperin JM, Wolf IE, Pascualvaca DM, Schmeidler J, O'Brien JD. The factor structure of ADHD items in DSM-II-R: internal consistency and

- external validation. *Journal of Abnormal Child Psychology* 1993; 21: 441–53.
- Hinshaw SP. On the distinction between attentional deficits/hyperactivity and conduct problem/aggression in child psychopathology. *Psychological Bulletin* 1987; 101: 443–63.
- Hobwenko H. *Attention Deficit/Hyperactivity Disorder*. London/Philadelphia: Jessica Kingsley Publishers, 1999.
- Holland ML, Gimpel GA, Merrel KW. Innovations in assessing ADHD: Development, psychometric properties, and factor structure of the ADHD Symptoms Rating Scale. *Journal of Psychopathology and Behavioural Assessment* 1998; 20: 307–32.
- Kalverboer AF. Disorders in attention and motor behaviour in children (De nieuwe buitenbeentjes. Stoornissen in aandacht en motoriek bij kinderen). Rotterdam: Lemniscaat, 1996.
- Lahey BA, Carlson CL. Validity of the diagnostic category of attention deficit disorder without hyperactivity: a review of the literature. *Journal of Learning Disabilities* 1991; 24: 110–20.
- Lahey BA, Pelham WE, Schaughency E, Atkins MS. Dimensions and types of attention deficit disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* 1988; 27: 330–5.
- Landis JR, Koch GG. The measurement of Observer Agreement for categorical Data. *Biometrics* 1977; 33: 159–74.
- Loeber R, Lahey BA, Thomas C. The diagnostic conundrum of oppositional defiant disorder, conduct disorder and its subtypes. *Journal of Abnormal Psychology* 1991; 100: 379–90.
- Long JS. *Confirmatory Factor Analysis*. London: Sage, 1983.
- Munden A, Arcelus J. *The ADHD-Handbook*. London/Philadelphia: Jessica Kingsley Publishers, 1999.
- Nelson-Gray RO. DSM-IV: Empirical guidelines from psychometrics. *Journal of Abnormal Psychology* 1991; 100: 308–15.
- McCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structural modelling. *Psychological Methods* 1996; 1: 130–49.
- Pelham WE, Gnacy EM, Greenslade KE, Milich R. Teacher ratings of DSM-III-R symptoms for the disruptive behaviour disorders. *J. Am. Acad. Adolesc. Psychiatry* 1992; 31: 210–18.
- Pillow DR, Pelham WE, Hoza B, Molina BSG, Stultz CH. Confirmatory factor analysis examining attention deficit hyperactivity disorder symptoms and other childhood disruptive behaviours. *J. Abn. Child Psych.* 1998; 26: 293–310.
- Quay HC. Classification. In HC Quay and JS Werry (eds). *Psychopathological Disorders of Childhood*. New-York: John Wiley, 1979.
- Sattora A, Bentler PM. Corrections to test statistics and standard errors in covariance structure analysis. In A Van Eye and CC Clogg (eds) *Latent Variables Analysis: Applications for Developmental Research*. Thousand Oaks: Sage, 1994.
- Swanson JM, Sergeant JA, Taylor E, Sonuga-Barke EJS, Jensen PS, Cantwell DP. Attention-deficit hyperactivity disorder and hyperkinetic disorder. *The Lancet* 1998; 351: 429–33.
- Scholte EM. Social Welfare in Holland. In M Colton, F Casas, M Drakeford, S Roberts, EM Scholte and M Williams (eds) *Stigma and Social Welfare: An International Comparative Study*. Aldershot: Avebury, 1997.
- Scholte EM, Van der Ploeg JD. Exploring factors governing successful residential treatment of youngsters with serious behavioural difficulties. Findings from a longitudinal study in Holland. *Childhood. A Global Journal of Child Research* 1000; 7: 129–53.
- Taylor, E. Attention Deficit and Conduct Disorder Syndromes. In M Rutter, M Hussian, A Tuma, IS Lann (eds) *Assessment and Diagnosis in Child Psychopathology*. London: David Fulton, 1988.
- Taylor E. Syndromes of attention Defiict and Overactivity. In M Rutter, E Taylor, L Hersov (eds) *Child and Adolescent Psychiatry. Modern Approaches (2 edn)*. Oxford: Blackwell, 1994.
- Van der Ploeg JD, Bruininks ACM (1996) *ADHD – Problem and Methods of Treatment*. Delft/Amsterdam: Eburon/Nippo, 1996.
- Wenar C, Kerig P. *Developmental Psychopathology. From Infancy through Adolescence*. New-York: McGraw-Hill, 1999.

#### Appendix: DSM-III and DSM-IV related items of the professional care workers' questionnaire

1. Fidgets with hands or feet or squirms in seat. (10)
2. Leaves set seat in situations in which remaining seated is expected. (11)
3. Is easily distracted by extraneous stimuli. (8)
4. Has difficulty awaiting turn. (17)
5. Blurts out answers before questions have been completed. (16)
6. Does not follow through on instructions, fails to finish schoolwork or duties. (4)
7. Makes careless mistakes in schoolwork, work, or other activities. (1)
8. Has difficulty in engaging in activities quietly. (13)
9. Interrupts or intrudes on others, butts into conversations or games. (18)
10. Does not seem to listen when spoken to. (3)
11. Loses things necessary for tasks or activities, for example toys, pencils, books. (7)

12. Undertakes dangerous activities, runs about or climbs excessively. (12)
13. Has difficulty organizing tasks and activities. (5)
14. Avoids engaging in tasks that require sustained mental effort. (6)
15. Acts as if 'driven by a motor'. (14)
16. Talks excessively. (15)
17. Has difficulty sustaining attention in tasks or play activities. (2)
18. Is forgetful in daily activities. (9)

*Correspondence: EM Scholte, Leiden University, Department of Education, PO Box 9555, 2300 RB Leiden, The Netherlands.*

(References to DSM-IV in brackets.)