<https://helda.helsinki.fi>

Retrospective Assessment of ADHD Symptoms in Childhood: Discriminatory Validity of Finnish Translation of the Wender Utah Rating Scale

Kivisaari, Sasa

2012

Kivisaari , S , Laasonen , M , Leppämäki , S , Tani , P & Hokkanen , L 2012 , ' Retrospective Assessment of ADHD Symptoms in Childhood: Discriminatory Validity of Finnish Translation of the Wender Utah Rating Scale ' , Journal of Attention Disorders , vol. 16 , no. 6 , pp. 449-459 . <https://doi.org/10.1177/1087054710397801>

<http://hdl.handle.net/10138/252770>

<https://doi.org/10.1177/1087054710397801>

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

Retrospective Assessment of ADHD Symptoms in Childhood: Discriminatory Validity of Finnish Translation of the Wender Utah Rating Scale

Journal of Attention Disorders
16(6) 449–459
© 2012 SAGE Publications
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1087054710397801
<http://jad.sagepub.com>

Sasa Kivisaari¹, Marja Laasonen¹, Sami Leppämäki¹,
Pekka Tani¹, and Laura Hokkanen¹

Abstract

Objective: To examine the discriminatory validity of the Wender Utah Rating Scale (WURS) and its five suggested subscales (Conduct Problems, Impulsivity Problems, Mood Difficulties, Inattention/Anxiety, Academic Concerns) in a Finnish sample. **Method:** WURS was administered to 114 adults, aged 18 to 55 years. Participants with ADHD ($n = 37$) and dyslexia ($n = 36$) were compared with healthy controls ($n = 41$). **Results:** The ADHD group scored significantly higher than the control group on all subscales. Compared with the dyslexia group, the ADHD group did not differ in Mood Difficulties or Academic Concerns. Using the total score, the positive predictive value was .53 in this sample and only .21 when the prevalence of ADHD was taken into account. **Conclusion:** Three out of five domains of WURS are reliable indicators of ADHD. Domains with low discriminatory power, low general prevalence of ADHD, and other developmental disorders within the population decrease the accuracy. (*J. of Att. Dis.* 2012; 16(6) 449-459)

Keywords

ADHD, Wender Utah Rating Scale, self-rating scales, dyslexia, DyAdd

ADHD was historically considered a childhood disorder where the symptoms would substantially decline or completely remit with time. Although the persistence of the syndrome has since been acknowledged, the diagnosis of adult ADHD remains a challenge. In the current diagnostic criteria, the symptom onset has to be retrospectively dated to childhood (American Psychiatric Association, 2000; World Health Organization, 1998), which can be difficult if reliable informants or appropriate school records are unavailable. As a result, the diagnosis of adults relies strongly on the information provided by the individual being assessed. Recently, the requirement of proven impairment before the age of 7 has been challenged and modifications to better address the issue in the upcoming *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; *DSM-V*; American Psychiatric Association, 2000) criteria have been suggested (Bell, 2010). Still, ADHD is a developmental condition, and the quality and quantity of symptoms in childhood need to be assessed. In clinical practice, behavioral self-rating scales are essential as they enable the subjective complaints to be compared with a normative standard. Wender Utah Rating Scale (WURS) is one such widely used clinical method developed for retrospectively surveying childhood

symptoms of ADHD (Ward, Wender, & Reimherr, 1993). As a part of the Project Adult Dyslexia and Attention Deficit Disorder (DyAdd; Laasonen, Leppämäki, Tani, & Hokkanen, 2009), we investigated the discriminatory validity of the WURS (25-item short form) in a Finnish sample of adults, focusing on dyslexia as a confounding factor.

The basic symptom domains of ADHD, as defined by the diagnostic classification systems, are hyperactivity, impulsivity, and inattention. The behavioral manifestations of these symptoms comprise the clinical picture of ADHD, on which the diagnosis is currently solely based (American Psychiatric Association, 2000; World Health Organization, 1998). Empirically, ADHD is not limited to the basic symptoms but can be associated with cognitive deficits (Frazier, Demaree, & Youngstrom, 2004; Hervey, Epstein, & Curry, 2004) and functional impairment in various life domains (e.g., Biederman et al., 1998, 2006). For example, ADHD is

¹University of Helsinki, Helsinki, Finland

Corresponding Author:

Laura Hokkanen, Institute of Behavioural Sciences, Division of Cognitive and Neuropsychology, University of Helsinki, Finland
Email: laura.hokkanen@helsinki.fi

related to academic underachievement (Biederman et al., 2008; Roy-Byrne et al., 1997), and individuals with ADHD often have a history of school problems that cannot be explained by learning disabilities (Seidman, Biederman, Weber, Hatch, & Faraone, 1998).

For the purpose of clinical needs, the self-rating scales often aim to cover more of the ADHD syndrome than is required for the formal diagnostic criteria. The WURS is based on the Utah criteria of ADHD (Wender, 1995), but it also includes items that tap nondiagnostic symptoms, particularly those of negative mood and affective lability (Fossati et al., 2001; McCann, Scheele, Ward, & Roy-Byrne, 2000; Norvilitis, Ingersoll, Zhang, & Jia, 2008; Oncu, Olmez, & Senturk, 2005; Retz-Junginger et al., 2003; Suhr, Zimak, Buelow, & Fox, 2009), educational difficulties (Norvilitis et al., 2008; Oncu et al., 2005; Suhr et al., 2009), and conduct problems (McCann et al., 2000; Norvilitis et al., 2008). Nondiagnostic items have clinical utility, but at the same time, they may reduce the diagnostic specificity of the instrument, as patients with other psychiatric, neuropsychiatric, and developmental disorders might systematically gain high scores. Elevated scores on the WURS have been reported in personality disorders (Fossati, Novella, Donati, Donini, & Maffei, 2002; Lindberg et al., 2004) and affective disorders, such as major depression and bipolar disorder (Oncu et al., 2005).

One of the challenges in diagnostic decision making is that different disorders truly have overlapping symptomatology. A common developmental condition that often coexists with ADHD is dyslexia: It has been estimated that 9% to 45% of the individuals with either ADHD or dyslexia also meet the criteria for the other disorder (Carroll, Maughan, Goodman, & Meltzer, 2005; Dykman & Ackerman, 1991; Willcutt & Pennington, 2000). Even when full criteria for both disorders are not met, the symptoms can overlap. Individuals diagnosed with dyslexia can manifest elevated levels of the core symptoms of ADHD (Carroll et al., 2005), experience cognitive problems (Brambati et al., 2006; Carroll et al., 2005), and by definition they show educational underperformance (Snowling, Muter, & Carroll, 2007). Reasons behind the underperformance probably vary in dyslexia and ADHD and need to be analyzed separately; the existence of the underperformance symptom alone may not differentiate between the two groups. To examine the discriminatory ability of the WURS, the self-ratings of individuals with ADHD are, in this study, compared with not only healthy controls but also with individuals with dyslexia. To the best of our knowledge, this comparison has not yet been conducted with the WURS.

The diagnostic properties of the WURS have typically been studied by assigning a cutoff for the total score and reporting diagnostic classification statistics, such as sensitivity and specificity for discriminating cases with ADHD from those without. Two major problems have emerged. First, except for the original validation study (Ward et al., 1993),

the total score has proven limited in its specificity in predicting ADHD diagnosis, especially with relation to clinical control groups (Oncu et al., 2005; Roy-Byrne et al., 1997; Suhr et al., 2009). The approach of the present article is to examine the WURS through its subscales in addition to the total score. The underlying assumption is that some item clusters of the WURS might show better discriminatory abilities than others. The subscales in this study are adopted from Suhr et al. (2009): Conduct Problems, Impulsivity Problems, Mood Difficulties, Inattention/Anxiety Symptoms, and Academic Concerns. Alternative factor structures have also been reported (see Caci, Bouchez, & Baylé, *in press*), but this was chosen because of the larger sample size of the original study and because the study included control groups with psychological symptoms or diagnoses, in addition to ADHD.

Second problem in previous studies is that sensitivity and specificity statistics are of limited use for a clinician as the statistics are always bound to the sample and do not provide direct implications for a given test result in a different setting. The probability of a correct classification can be reported in the form of positive and negative predictive values, but these statistics are poorly generalizable outside the derivation sample, especially if the prevalence of the disorder in the general population is low (the so-called base-rate problem; see Reid & Maag, 1994). If the prevalence of the disorder in a given population is known, the problem can be overcome by mathematical adjustments (Bergman & Pantell, 1984), which is the approach in the present study.

To summarize, the aims of this study are twofold. First, to evaluate the discriminatory ability of the WURS by comparing the scores in the subscales of the WURS in three groups (ADHD, dyslexia, and controls) and second, to assess the diagnostic accuracy of the total score in the present sample and in general population, corrected for the low prevalence of the disorder.

Method

This study was carried out in the Department of Psychology at the University of Helsinki, in collaboration with the Helsinki University Central Hospital, and it was approved by the Ethics Committee of the Hospital District of Helsinki and Uusimaa. As described elsewhere (Laasonen et al., 2009), a wide array of neuropsychological tests and experimental methods were administered to adults with ADHD and/or dyslexia who volunteered to participate in the project DyAdd. The WURS self-rating questionnaire was sent to all participants who met the inclusion, but not the exclusion, criteria.

Participants

The self-rating questionnaire was administered to 123 participants between the ages of 18 and 55 years falling into

three categories of equal size: individuals with ADHD, individuals with dyslexia, and healthy controls. An informed consent was acquired from each participant. All participants confirmed the lack of current or previous history of psychiatric or neurological conditions and substance abuse.

The ADHD group originally consisted of 41 adults, the majority of whom had been recruited from the neuropsychiatric outpatient clinic of the Helsinki University Central Hospital. A smaller proportion was recruited from private practices.

ADHD was diagnosed according to the *DSM-IV* (American Psychiatric Association, 2000) criteria using *Conners' Adult ADHD Diagnostic Interview for DSM-IV* (Epstein, Johnson, & Conners, 2001) by a medical doctor specialized in neuropsychiatry (authors Sami Leppämäki or Pekka Tani, in most cases). Thus, hyperactivity was not a required characteristic and those with inattention only were included. Multiple sources of information (interview of patient's parents when possible, all relevant school and medical records) were used to confirm childhood symptoms. Confounding psychiatric disorders were excluded by the *Structured Clinical Interview for DSM-IV Axis I Disorders* (SCID-I; First, Spitzer, Gibbon, & Williams, 1996) and SCID-II interviews (First, Gibbon, Spitzer, Williams, & Benjamin, 1997). WURS was not used in the diagnostic process.

Of the 41 participants in the ADHD group, 9 (22%) were diagnosed with dyslexia (the comorbid group), and these participants were excluded from the discriminatory analyses. Those with ADHD participated in the project unmedicated.

The dyslexia group comprised 41 adults who reported having a dyslexia diagnosis but who did not have a diagnosis of ADHD. One of the participants from the dyslexia group was excluded on the basis of lowered full scale intelligence quotient (FSIQ, International Classification of Diseases [ICD-10] criteria; World Health Organization, 1998), after which the size of the group was 40. The majority of the participants with dyslexia were recruited through HERO (diverse learners' association in Helsinki). Their diagnosis was based on achievement criteria that varied slightly.

Therefore, the current phonological processing and reading status of each participant in the dyslexia and comorbid groups was checked against the age-corrected values of our previous (Laasonen, 2002) and current control data. Participants in these two groups performed below -1 standard deviation in phonological processing and reading, as assessed with phonological naming speed (rapid alternate stimulus naming speed/accuracy; Wolf, 1986), phonological awareness (phonological synthesis accuracy; Laasonen, Service, & Virsu, 2002), phonological memory (Wechsler Adult Intelligence Scale [WAIS] digit span forward length; Wechsler, 2005), and reading (oral reading speed/accuracy; task details in Laasonen et al., 2002).

The Control group consisted of 41 volunteers. The participants in this group confirmed the lack of history of learning deficits or ADHD. The control group and a proportion of the dyslexia group were recruited by advertisements.

Measures

The full neuropsychological assessment battery of the Project DyAdd is presented elsewhere (Laasonen et al., 2009). In all the assessments, the examiner was blind to the group of the participant. The general performance level was estimated using four subtests of the WAIS-III (Wechsler, 2005): vocabulary, similarities, matrix reasoning, and block design. The WURS was mailed to the participants beforehand and returned to the examiner at the assessment. If there were any questions regarding the questionnaire, they were discussed at this point.

In the WURS, participants are instructed to rate items on the basis of recollections of childhood behavior. Each item is measured on a 5-point Likert-type scale (0 = *not at all or very slightly*, 1 = *mildly*, 2 = *moderately*, 3 = *quite a bit*, 4 = *very much*). The items are summed to constitute the total score which can vary from 0 to 100. In this study, the participants filled the original WURS questionnaire with 61 items, but only the 25 items of the short version (Ward et al., 1993) were analyzed as they are the most commonly used items in the literature. The Finnish version of the WURS was translated by an author of this study (Pekka Tani), after which it was back-translated by a professional translator.

The five subscales used in this study were chosen from Suhr et al. (2009), where they were formed on the basis of the principal component analysis of a large sample ($n = 1431$). The subscales, the items constituting each subscale, and their Cronbach's alphas are shown in Table 1. The internal consistencies were adequate in all but the Academic Concerns subscale where it was low ($\alpha < .60$), at least partly reflecting the small number of items.

Statistical Analyses

All statistical analyses were performed using SPSS 18.0 for Windows software (SPSS Inc., USA).

Six out of 122 participants (4.9%) failed to return the questionnaire or did not provide sufficient demographic information. Three of the participants were from the dyslexia group and three from the ADHD group. In addition, the questionnaires of two participants (from ADHD and dyslexia groups) were removed from the data because of too many missing values ($>12\%$). Thus, the size of the final data was 114 (mean age = 35.2 years, $SD = 10.7$ years; 57 males; mean years of education = 13.3, $SD = 3.12$ years). The remaining missing values were tested by chi-square test and were found unassociated with group membership or any of the demographic variables. Therefore, they were replaced

Table 1. The Items Constituting the Five Subscales of the WURS^a and the Cronbach's Alphas of Each

Conduct Problems	Impulsivity Problems	Mood Difficulties	Inattention/Anxiety	Academic Concerns
Hot-/short tempered, low boiling point	Trouble with stick- to-it-iveness, not following through	Sad or blue, depressed, unhappy	Concentration problems, easily distracted	Overall a poor student, slow learner
Temper outbursts, tantrums	Acting without thinking, impulsive	Low opinion of self	Anxious, worrying	Trouble with math/ numbers
Stubborn, strong willed	Tend to be immature	Guilty, regretful	Nervous, fidgety	Did not achieve up to potential
Disobedient with parents, rebellious	Lose control of self	Unpopular with other children, did not keep friends for long	Inattentive, daydreaming	
Irritable	Tend to be or act irrational			
Moody, have ups and downs	Trouble seeing things from other's view			
Angry				
Trouble with authority				
$\alpha = .862$	$\alpha = .778$	$\alpha = .755$	$\alpha = .821$	$\alpha = .579$

^aSuhr, Zimak, Buelow, and Fox, 2009.

by maximum likelihood-based expectation and maximization method.

Six out of 122 participants (4.9%) failed to return the questionnaire or did not provide sufficient demographic information. Three of the participants were from the dyslexia group and three from the ADHD group. In addition, the questionnaires of two participants (from ADHD and dyslexia groups) were removed from the data because of too many missing values (>12%). Thus, the size of the final data was 114 (mean age = 35.2 years, $SD = 10.7$ years; 57 males; mean years of education = 13.3, $SD = 3.12$ years). The remaining missing values were tested by chi-square test and were found unassociated with group membership or any of the demographic variables. Therefore, they were replaced by maximum likelihood-based expectation and maximization method.

The total scores were examined by receiver operating characteristic (ROC) curves, and the cutoff scores were chosen on the basis of the total correct classifications and Youden's J (sensitivity + specificity - 1). First, the control and the dyslexia group pooled together formed the reference (Situation A). Second, the dyslexia group was excluded and only the control group was considered as the reference (Situation B). The comorbid group was excluded from both situations. Situation B was used further in dealing with the base-rate problem. Here, the positive and negative predictive values were recalculated using the formula that includes the prevalence,¹ and the sample-based figures were substituted by the known prevalence of the condition (Bergman & Pantell, 1984). The prevalence of adult ADHD was set to 4.4% (Kessler et al., 2006).

Results

The groups did not differ in the mean age, the education level, or the male/female ratios (Table 2). A significant main effect was found for the WAIS FSIQ.

Group Comparisons and Discriminatory Power of the Subscales

In a MANOVA, the WURS total score was found to differ between the three groups, $F(2, 111) = 39.3, p < .0001, \eta^2 = .41$. A 3×5 MANOVA revealed a significant main effect of group also for the subscales of the WURS, Wilks's $\Lambda = .40$, $F(10, 214) = 12.4, p < .0001$, partial $\eta^2 = .37$. Significant main effects of group were found for all the subscales: Conduct Problems, $F(2, 111) = 21.8, p < .0001, \eta^2 = .28$; Impulsivity Problems, $F(2, 111) = 28.2, p < .0001, \eta^2 = .34$; Mood Difficulties, $F(2, 111) = 6.26, p = .003, \eta^2 = .10$; Inattention/Anxiety, $F(2, 111) = 26.3, p < .0001, \eta^2 = .32$; and Academic Concerns, $F(2, 111) = 36.6, p < .0001, \eta^2 = .40$. The pairwise Bonferroni corrected comparisons of the subscales as well as the WURS total score are shown in Table 3.

To further elaborate the group differences, the subscale scores of the WURS were entered at the same time as predictors into a discriminant analysis. The comorbid group was excluded from the analysis to enhance the differences between the groups. Two discriminant functions were extracted and their combination significantly predicted group membership, $\Lambda = .37, \chi^2(10) = 100.8, p < .001$. After the removal of the first function, the second remained statistically significant, $\Lambda = .78, \chi^2(4) = 24.4, p < .001$, canonical $r = .47$.

Table 2. Demographic Data for the ADHD, Control, and Dyslexia groups ($N = 114$)

	ADHD ($n = 37$)	Dyslexia ($n = 36$)	Control ($n = 41$)	F / χ^2	p
Gender, n (%)				1.7 ^a	.43
Female	16 (43.2)	21 (58.3)	21 (51.2)		
Male	21 (56.8)	15 (41.7)	20 (48.8)		
Education, n (%)				6.4 ^a	.17
Basic	23 (62.2)	18 (50.0)	15 (36.6)		
Middle	6 (16.2)	11 (30.6)	12 (29.3)		
High	8 (21.6)	7 (19.4)	14 (34.1)		
Age, M (SD)	32.0 (8.8)	35.8 (10.7)	37.5 (11.8)	2.8 ^b	.07
WAIS IQ ^c , M (SD)	103.7 (14.8)	106.2 (11.7)	113.3 (11.3)	5.9 ^b	.004

Note: WAIS = Wechsler Adult Intelligence Scale.

^aChi-square test.

^bOne-way ANOVA.

^cIQ estimation based on vocabulary, similarities, matrix reasoning, and block design.

Table 3. Pairwise Group Comparisons ($N = 114$)

	ADHD ($n = 37$)	Dyslexia ($n = 36$)	Control ($n = 41$)	ADHD versus control (p value)	ADHD versus dyslexia (p value)	Control versus dyslexia (p value)
WURS total	50.5 (13.9)	37.0 (14.3)	22.9 (12.9)	<.001	<.001	<.001
Conduct Problems	16.7 (6.2)	10.4 (6.1)	8.1 (5.5)	<.001	<.001	.28
Impulsivity Problems	11.9 (4.2)	8.0 (4.3)	5.0 (3.7)	<.001	<.001	.004
Mood Difficulties	6.7 (3.5)	5.5 (3.6)	4.0 (3.1)	.002	.40	.17
Inattention/Anxiety	9.0 (3.4)	6.6 (3.5)	3.7 (2.9)	<.001	.01	.001
Academic Concerns	6.1 (2.5)	6.5 (2.7)	2.2 (2.2)	<.001	1.0	<.001

Note: WURS = Wender Utah Rating Scale. The table shows the mean scores by group, standard deviations as well as Bonferroni corrected p values for differences.

The results of the discriminant analysis were in line with the MANOVA results. The first discriminant function distinguished the control group from the clinical groups, most notably from the ADHD group (Figure 1). All structure coefficients (Table 4) were positive, Inattention/Anxiety having the highest and Mood Difficulties the lowest loading. In the second discriminant function, the dyslexia group was separated from the ADHD and the control groups. The Academic Concerns score had the largest negative and the Conduct Problems the largest positive loading.

Diagnostic Accuracy

The calculations of diagnostic accuracy were first conducted with the control and the dyslexia groups pooled together (Situation A; Table 5). The optimal cutoff score in this sample was 37 (area = .84, 95% confidence interval [CI] = [.77-.91]). The commonly used cutoff score of 46 in Situation A yielded a higher positive predictive value but a lower sensitivity.

To calculate the discriminatory power when no clinical controls are included, a second analysis was conducted and

the ROC curve was now examined using only the control group as a reference (Situation B; Table 5). With the dyslexia group excluded, the general accuracy improved (area = .93, 95% CI = [.87-.98]) and the optimal cutoff score was now 36. To deal with the base-rate problem, the results of this analysis were then recalculated using the estimated prevalence of ADHD, and the probability of correct positive classification (positive predictive value for the cutoff score 36) fell down to .21. The probability of a correct negative classification (negative predictive value) was .99, illustrating the low prevalence. The positive predictive value for the cutoff score 46, when adjusted for the known prevalence of ADHD, was .60.

Discussion

There is a demand for empirically validated assessment methods for adult ADHD to accompany the formal diagnostic criteria, and the general aim of this study was to examine the discriminatory validity of the WURS. We wanted to explore this widely used self-rating scale, especially in relation to dyslexia, which is a common comorbid condition to

Figure 1. Group centroids (squares) and observed discriminant function scores of the individual participants by group. The centroids represent mean discriminant function scores for each group on a given discriminant function.

Table 4. Group Centroids and Structure Coefficients for Two Discriminant Functions Derived From the Subscale Scores of the WURS

Function	Centroids			Structure coefficients				
	ADHD	Dyslexia	Control	Conduct Problems	Impulsivity Problems	Mood Difficulties	Inattention/Anxiety	Academic Concerns
1	1.30	0.41	-1.25	0.54	0.63	0.33	0.74	0.68
2	0.58	-0.69	0.21	0.56	0.35	0.11	0.33	-0.66

ADHD. In the sample pool of ADHD individuals referred to us, 22% had both conditions. The discriminatory properties of the WURS were examined through its suggested subscales: Conduct Problems, Impulsivity Problems, Mood Difficulties, Inattention/Anxiety Symptoms, and Academic Concerns (Suhr et al., 2009). The approach enabled a more in-depth examination of the characteristics of the multidimensional rating scale.

The ADHD group gained significantly higher scores than the control group on all five subscales of the WURS. We even included a comorbid group with both ADHD and

dyslexia in our initial analyses to see if the symptoms of ADHD would still be adequately distinguishable, and it seems they were. The group size was relatively small and decreased further when the comorbid group was excluded, but the participants were carefully diagnosed. The discriminant analysis revealed that the two subscales describing the basic symptoms of ADHD (Inattention/Anxiety and Impulsivity Problems) together with Academic Concerns had the most discriminatory power with respect to the healthy controls. However, when the ADHD group was compared with the dyslexia group, the mean scores of Mood

Table 5. Diagnostic Accuracy Using Different Cutoff Scores

	Situation A (total <i>n</i> = 105)				Situation B (total <i>n</i> = 69)			
	36	37 ^a	38	46	35	36 ^a	37	46
Sensitivity	.89	.86 ^a	.79	.64	.89	.89 ^a	.86	.64
Specificity	.69	.73 ^a	.73	.86	.83	.85 ^a	.88	.98
Positive predictive value	.51	.53 ^a	.51	.62	.78	.81 ^a	.83	.95
Negative predictive value	.95	.93 ^a	.90	.87	.92	.92 ^a	.90	.80
Total correct classifications	.74	.76 ^a	.74	.80	.85	.87 ^a	.87	.84
Youden's J	.58	.58 ^a	.51	.50	.72	.75 ^a	.74	.62
Kappa	.47	.49 ^a	.44	.49	.70	.73 ^a	.73	.65

Note: In Situation A, the control and dyslexia groups are pooled together. In Situation B, the dyslexia group is excluded. In both situations, the comorbid subgroup was excluded from the ADHD group.

^aOptimal cutoffs.

Difficulties and Academic Concerns were not significantly different, which suggests that these subscales have limited value in discriminating between the two clinical conditions. Both these subscales consist of only a small number of items, so their impact on the total score is relatively low. It is noteworthy but not surprising that the Academic Concerns subscale was more strongly associated with dyslexia than ADHD. The finding that the Mood Difficulties subscale differentiated poorly between the groups was consistent with previous results. In the study by McCann et al. (2000), this domain did not reliably distinguish the ADHD group from the nonclinical control participants. Similarly, Oncu et al. (2005) found that the factor scores of their depression domain did not significantly differ between participants with ADHD, depression, and bipolar disorder.

The diagnostic accuracy of the WURS was first examined using the controls and dyslexia group pooled together as the reference group. With a cutoff score of 37, the proportion of correct positive classifications for ADHD was 53%, the proportion of correct negative classifications was 93%, and that of total correct classifications 76%. The more commonly used cutoff point of 46 (Ward et al., 1993) was able to correctly classify 80% of all cases, but the sensitivity was lower, and more than a third of those with ADHD scored below this cutoff point. Partially, different results may be explained by the different diagnostic criteria used: In the *DSM-IV*, hyperactivity is not a required symptom, whereas in the Utah criteria it is. The trade-off between sensitivity and specificity in the present sample suggests using a lower cutoff point than previously recommended, and similar results have been obtained by others (McCann et al., 2000; Oncu et al., 2005; Roy-Byrne et al., 1997; Suhr et al., 2009). High specificity often comes with the price of low sensitivity and vice versa, and the users of the rating scale should be aware of this caveat. Furthermore, these results are reliably generalizable only into a context where the estimated proportion of healthy individuals and those with ADHD

and other similar disorders resembles the current sample (i.e., one third of healthy normal individuals, one third of individuals with ADHD, and one third of individuals with other learning disabilities). This could correspond to a situation, for instance, in neuropsychiatric units, where differential diagnostics is commonplace.

To elucidate a situation outside the clinical context, such as screening in the general population, another set of calculations were conducted. When only controls and the ADHD group were included in the analysis, and the dyslexia group was excluded, the probability of total correct classifications with a cutoff score of 36 increased to 87%, but this value is overly optimistic as it would require a situation where all other clinical conditions such as learning disorders are non-existent, and ADHD represents almost half of the population. Even in this analysis, the high ratios of sensitivity (96%) and specificity (96%) reported by Ward et al. (1993) were not observed. To estimate the accuracy in the general population, the prevalence of adult ADHD was set to 4.4% (Kessler et al., 2006). When this approach was selected, the probability of a correct positive classification fell from 81% to 21%. The negative predictive value was .99, but this only reflects the low base rate of the syndrome—if there are very few cases in the population, the likelihood of them being mistakenly picked up is low. Raising the cutoff score increases the probability of the correct positive classifications, and with the commonly used score of 46, the probability of correct diagnosis approached 60%. The exact figure depends on the prevalence and in adults the estimations vary, but the American population-based epidemiological survey data (Kessler et al., 2006) are used here to illustrate the point. Epidemiological studies of ADHD indicate a prevalence rate of 6.6% to 8.5% in Finnish children and adolescents (Puura et al., 1998; Smalley et al., 2007), so we believe the figures are of the right magnitude. Although the procedure we adopted simplifies the real-life situation by ignoring comorbidity, the results show that in populations with a low

prevalence rate, the risk of false positive classifications increases considerably.

The dyslexia group had consistently higher scores on the WURS than the control group, which corresponds to the studies that have demonstrated the overlapping symptomatology of dyslexia and ADHD (Carroll et al., 2005; Dykman & Ackerman, 1991; Willcutt & Pennington, 2000). Particularly frequent among the individuals with dyslexia were understandably the Academic Concerns, but elevated scores were observed also on Impulsivity Problems and Inattention/Anxiety. As a consequence, a substantial amount of the participants with dyslexia were classified into the ADHD group when using a simple, total score cutoff. This is a problem in terms of diagnostic accuracy but can be understood in light of recent findings of shared cognitive features in ADHD and dyslexia. Based on new evidence from a large number of studies, Shaywitz and Shaywitz (2008) have argued that the generation of the phonological code from print that is demanded in reading is not automatic in dyslexic readers but requires attention and that a disruption of attentional mechanisms plays a causal role in reading difficulties. In previous articles of the Project DyAdd, we have explored the similarities of the cognitive profiles in our participants with dyslexia, ADHD, and their comorbid combination (Laasonen et al., 2009) and also examined the phonological processing and academic skills in these groups (Laasonen, Lehtinen, Leppämäki, Tani, & Hokkanen, 2010). In WAIS-III (Wechsler, 2005), FSIQ and the processing-speed factor within it were found to be impaired for all clinical groups. A specific difficulty in performance IQ emerged in those with ADHD, whereas only those with dyslexia were impaired in the working-memory factor. The comorbid group evidenced no specific difficulties (Laasonen et al., 2009). In the domain of phonological processing, the ADHD group did not show impairments, although the accuracy of technical reading, spelling, and arithmetic posed difficulties to them. However, these difficulties disappeared when IQ was taken into consideration, and thus the academic difficulties were perhaps more attributable to slightly lowered general ability or difficulties in attention and executive functions (Laasonen et al., 2010). Taken together, all these findings point toward attentional mechanisms playing a part both in reading and with other academic skills in dyslexia and ADHD, which can explain overlapping results also on a rating scale exploring school performance.

The assignment of a cutoff score is always arbitrary to some degree, and a mechanical use of a cutoff will almost inevitably lead to misclassifications (Reid & Maag, 1994). In consequence, one should not focus only on the cutoff score but also on the way in which scores are distributed along the rating scale. The results of the present study do not indicate that retrospective assessments are unreliable; rather, they stress the fact that some item domains of the WURS are not specific to ADHD, especially when compared

with a clinical-control group. In clinical practice, high ratings on either the Mood Difficulties or Academic Concerns subscale warrant further assessment of other conditions that might affect optimal academic achievement, including developmental and learning disabilities, and additional assessment methods should be used for identifying the primary deficit. However, given that the WURS is sometimes employed in research purposes to retrospectively study the association between current psychopathology and childhood ADHD (e.g., Fossati et al., 2002; Lindberg et al., 2004), the findings of this study combined with previous results (McCann et al., 2000; Oncu et al., 2005; Roy-Byrne et al., 1997; Suhr et al., 2009) are noteworthy. It is questionable whether the WURS alone can make a reliable distinction between multiple types of disorders. To illustrate, the emphasis on affective symptoms of ADHD and, at the same time, the poor discriminatory power of the Mood Difficulties subscale make it possible that symptoms of any disorder having adverse effect on mood can be misinterpreted as indicators of ADHD.

A potential limitation to the current study is that all participants in the ADHD group, but not in the other two groups, had recently been referred to a specialist to formulate the diagnosis, and during the process, they had confirmed the childhood and current symptoms of ADHD by the *DSM-IV* criteria. Hence, the participants in this study may have been more prone to report their ADHD symptoms (compare with Adler et al., 2008) or able to describe their symptoms more accurately than individuals who have not been interviewed for a diagnosis. Therefore, even though the WURS specifically had not been administered, the study could have involved some degree of circularity as the respondents knew what types of answers were expected. Using a longer version of the WURS, but only analyzing the 25 commonly used items, helps in alleviating this problem somewhat: It is less clear to the respondents which are the items they are assumed to find difficulty with. This is a common problem everywhere as the Internet together with other media make information regarding ADHD diagnosis readily available for anyone seeking the diagnosis. In the current study, if anything, it should have improved the discrimination between the groups.

Another concern is the reliability of the answers. The mode of questionnaire administration can have an effect on the data acquired. We mailed the questionnaires to the participants and made sure they returned them when they came to the assessment and also answered any questions they might have at that point. According to Bowling (2005), this type of self-administration avoids some of the pitfall of interviewer and social-desirability bias but is prone to misunderstandings and item nonresponse. The missing values in our data were, however, unassociated with the group (ADHD or dyslexia), so this particular effect is unlikely to distort the results.

Conclusion

The results of this study show that the total score of the WURS alone is not a reliable predictor of childhood history of ADHD characteristics especially when used in populations with a low prevalence of ADHD or a high prevalence of other learning disorders. In comparison with the unidimensional total score, the subscales better reflect the multifaceted syndrome and can yield more detailed information about the presence and quality of ADHD symptoms. However, in diagnostic decision making as well as in research designs, it should be noted that items in Mood Difficulties and the Academic Concerns subscales were not specific to ADHD.

Declaration of Conflicting Interests

The author(s) declared that they had no conflicts of interests with respect to their authorship or the publication of this article.

Funding

Financial support was received from the Academy of Finland (projects 108410, 217065, and 217998), Emil Aaltonen Foundation, and Otologic Research Foundation.

Note

1. P (correct positive classification) = prevalence $\times$ sensitivity / ((prevalence $\times$ sensitivity) + (1 – prevalence) $\times$ (1 – specificity))
 P (correct negative classification) = (1 – prevalence) $\times$ specificity / ((1 – prevalence) $\times$ specificity + prevalence $\times$ (1 – sensitivity)).

References

- Adler, L. A., Faraone, S. V., Spencer, T. J., Michelson, D., Reimherr, F. W., Glatt, S. J., . . . Biederman, J. (2008). The reliability and validity of self- and investigator ratings of ADHD in adults. *Journal of Attention Disorders, 11*, 711-719.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual for mental disorder* (4th ed., text rev.). Washington, DC: Author.
- Bell, A. S. (2011). A critical review of ADHD diagnostic criteria: What to address in the DSM-V. *Journal of Attention Disorders, 15*, 3-10.
- Bergman, D. A., & Pantell, R. H. (1984). The art and science of medical decision making. *Journal of Pediatrics, 104*, 649-656.
- Biederman, J., Faraone, S. V., Spencer, T. J., Mick, E., Monuteaux, M. C., & Aleardi, M. (2006). Functional impairments in adults with self-reports of diagnosed ADHD: A controlled study of 1001 adults in the community. *Journal of Clinical Psychiatry, 67*, 524-540.
- Biederman, J., Faraone, S. V., Taylor, A., Sienna, M., Williamson, S., & Fine, C. (1998). Diagnostic continuity between child and adolescent ADHD: Findings from a longitudinal clinical sample. *Journal of the American Academy of Child & Adolescent Psychiatry, 37*, 305-313.
- Biederman, J., Petty, C. R., Fried, R., Kaiser, R., Dolan, C. R., Schoenfeld, S., . . . Faraone, S. V. (2008). Educational and occupational underattainment in adults with attention-deficit/hyperactivity disorder: A controlled study. *Journal of Clinical Psychiatry, 69*, 1217-1222.
- Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health, 27*, 281-291.
- Brambati, S. M., Termine, C., Ruffino, M., Danna, M., Lanzi, G., Stella, G., . . . Perani, D. (2006). Neuropsychological deficits and neural dysfunction in familial dyslexia. *Brain Research, 1113*, 174-185.
- Caci, H. M., Bouchez, J., & Baylé, F. J. (2010). An aid for diagnosing attention-deficit/hyperactivity disorder at adulthood: Psychometric properties of the French versions of two Wender Utah Rating Scales (WURS-25 and WURS-K). *Comprehensive Psychiatry, 51*, 325-331.
- Carroll, J. M., Maughan, B., Goodman, R., & Meltzer, H. (2005). Literacy difficulties and psychiatric disorders: Evidence for comorbidity. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 46*, 524-532.
- Dykman, R. A., & Ackerman, P. T. (1991). Attention deficit disorder and specific reading disability: Separate but often overlapping disorders. *Journal of Learning Disabilities, 24*, 96-103.
- Epstein, J. N., Johnson, D., & Conners, C. K. (2001). *Conners' adult ADHD diagnostic interview for DSM-IV (CAADID)*. Toronto, Ontario, Canada: Multi-Health Systems.
- First, M. B., Gibbon, M., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. S. (1997). *Structured clinical interview for DSM-IV Axis II personality disorders (SCID-II)*. Washington, DC: American Psychiatric Press.
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1996). *Structured clinical interview for DSM-IV Axis I disorders, clinician version (SCID-CV)*. Washington, DC: American Psychiatric Press.
- Fossati, A., Di Ceglie, A., Acquarini, E., Donati, D., Donini, M., Novella, L., & Maffei, C. (2001). The retrospective assessment of childhood attention deficit hyperactivity disorder in adults: Reliability and validity of the Italian version of the Wender Utah Rating Scale. *Comprehensive Psychiatry, 42*, 326-336.
- Fossati, A., Novella, L., Donati, D., Donini, M., & Maffei, C. (2002). History of childhood attention deficit/hyperactivity disorder symptoms and borderline personality disorder: A controlled study. *Comprehensive Psychiatry, 43*, 369-377.
- Frazier, T. W., Demaree, H. A., & Youngstrom, E. A. (2004). Meta-analysis of intellectual and neuropsychological test performance in attention-deficit/hyperactivity disorder. *Neuropsychology, 18*, 543-555.
- Hervey, A. S., Epstein, J. N., & Curry, J. F. (2004). Neuropsychology of adults with attention-deficit/hyperactivity disorder: A meta-analytic review. *Neuropsychology, 18*, 485-503.
- Kessler, R. C., Adler, L., Barkley, R., Biederman, J., Conners, C. K., Demler, O., . . . Zaslavsky, A. M. (2006). The prevalence and

- correlates of adult ADHD in the United States: Results from the national comorbidity survey replication. *American Journal of Psychiatry*, 163, 716-723.
- Laasonen, M. (2002). *Temporal acuity in developmental dyslexia across the life span: Tactile, auditory, visual, and crossmodal estimations*. Unpublished doctoral dissertation, University of Helsinki, Finland.
- Laasonen, M., Lehtinen, M., Leppämäki, S., Tani, P., & Hokkanen, L. (2010). Project DyAdd: Phonological processing, reading, spelling, and arithmetic in adults with dyslexia or ADHD. *Journal of Learning Disabilities*, 43, 3-14.
- Laasonen, M., Leppämäki, S., Tani, P., & Hokkanen, L. (2009). Adult dyslexia and attention deficit disorder in Finland—Project DyAdd: WAIS-III cognitive profiles. *Journal of Learning Disabilities*, 42, 511-527.
- Laasonen, M., Service, E., & Virsu, V. (2002). Crossmodal temporal order and processing acuity in developmentally dyslexic young adults. *Brain and Language*, 80, 340-354.
- Lindberg, N., Tani, P., Porkka-Heiskanen, T., Appelberg, B., Rimón, R., & Virkkunen, M. (2004). ADHD and sleep in homicidal men with antisocial personality disorder. *Neuropsychobiology*, 50, 41-47.
- McCann, B. S., Scheele, L., Ward, N., & Roy-Byrne, P. (2000). Discriminant validity of the Wender Utah Rating Scale for attention-deficit/hyperactivity disorder in adults. *Journal of Neuropsychiatry and Clinical Neurosciences*, 12, 240-245.
- Norvilitis, J. M., Ingersoll, T., Zhang, J., & Jia, S. (2008). Self-reported symptoms of ADHD among college students in China and the United States. *Journal of Attention Disorders*, 11, 558-567.
- Oncu, B., Olmez, S., & Senturk, V. (2005). Wender-Utah Derecelendirme Olcegi Turkce formunun eriskin dikkat eksikligi ve hiperaktivite bozuklugu'nda gecerlik ve guvenilirlik calismasi [Validity and reliability of the Turkish version of the Wender Utah Rating Scale for attention-deficit/hyperactivity disorder in adults.] *Turk Psikiyatri Dergisi*, 16, 252-259.
- Puura, K., Almqvist, F., Tamminen, T., Piha, J., Rasanen, E., Kumpulainen, K., . . . Koivisto, A. M. (1998). Psychiatric disturbances among prepubertal children in southern Finland. *Social Psychiatry and Psychiatric Epidemiology*, 33, 310-318.
- Reid, R., & Maag, J. W. (1994). How many fidgets in a pretty much: A critique of behavior rating scales for identifying students with ADHD. *Journal of School Psychology*, 32, 339-354.
- Retz-Junginger, P., Retz, W., Blocher, D., Stieglitz, R. D., Georg, T., Supprian, T., . . . Rösler, M. (2003). Reliabilität und Validität der Wender-Utah-Rating-Scale-Kurzform. Retrospektive Erfassung von Symptomen aus dem Spektrum der Aufmerksamkeitsdefizit/Hyperaktivitätsstörung [Reliability and validity of the Wender-Utah-Rating-Scale short form: Retrospective assessment of symptoms for attention deficit/hyperactivity disorder]. *Der Nervenarzt*, 74, 987-993.
- Roy-Byrne, P., Scheele, L., Brinkley, J., Ward, N., Wiatrak, C., Russo, J., . . . Varley, C. (1997). Adult attention-deficit hyperactivity disorder: Assessment guidelines based on clinical presentation to a specialty clinic. *Comprehensive Psychiatry*, 38, 133-140.
- Seidman, L. J., Biederman, J., Weber, W., Hatch, M., & Faraone, S. V. (1998). Neuropsychological function in adults with attention-deficit hyperactivity disorder. *Biological Psychiatry*, 44, 260-268.
- Shaywitz, S. E., & Shaywitz, B. A. (2008). Paying attention to reading: The neurobiology of reading and dyslexia. *Development and Psychopathology*, 20, 1329-1349.
- Smalley, S. L., McGough, J. J., Moilanen, I. K., Loo, S. K., Taanila, A., Ebeling, H., . . . Järvelin, M. R. (2007). Prevalence and psychiatric comorbidity of attention-deficit/hyperactivity disorder in an adolescent Finnish population. *Journal of the American Academy of Child & Adolescent Psychiatry*, 46, 1575-1583.
- Snowling, M. J., Muter, V., & Carroll, J. (2007). Children at family risk of dyslexia: A follow-up in early adolescence. *Journal of Child Psychology and Psychiatry*, 48, 609-618.
- Suhr, J., Zimak, E., Buelow, M., & Fox, L. (2009). Self-reported childhood attention-deficit/hyperactivity disorder symptoms are not specific to the disorder. *Comprehensive Psychiatry*, 50, 269-275.
- Ward, M. F., 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.
- Wechsler, D. (2005). *Wechsler Adult Intelligence Scale—Third Edition: Manual*. Helsinki, Finland: Psykologien Kustannus OY.
- Wender, P. (1995). *Attention-deficit hyperactivity disorder in adults*. New York, NY: Oxford University Press.
- Willcutt, E. G., & Pennington, B. F. (2000). Psychiatric comorbidity in children and adolescents with reading disability. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41, 1039-1048.
- Wolf, M. (1986). Rapid alternating stimulus naming in the developmental dyslexias. *Brain and Language*, 27, 360-379.
- World Health Organization. (1998). *The international statistical classification of diseases and related health problems, 10th revision* (2nd ed.). Helsinki, Finland: Author.

Bios

Sasa Kivisaari, MA, is a postgraduate student in the Institute of Behavioural Sciences, University of Helsinki.

Marja Laasonen, PhD, is a postdoctoral researcher in the Institute of Behavioural Sciences, Division of Cognitive and Neuropsychology at the University of Helsinki, and a neuropsychologist in the Department of Phoniatrics, Helsinki University Central Hospital, Helsinki, Finland. Her research interests include dyslexia, ADHD, and orofacial clefts.

Sami Leppämäki, MD, PhD, is a deputy chief physician at the Clinic for Neuropsychiatry, Department of Psychiatry, Helsinki University Central Hospital, and an adjunct professor of psychiatry

in the Department of Psychiatry, University of Helsinki, Finland. His research interests include affective disorders and developmental neuropsychiatric disorders.

Pekka Tani, MD, PhD, is a deputy chief physician at the Clinic for Neuropsychiatry, Department of Psychiatry, Helsinki University Central Hospital, and an adjunct professor of psychiatry in the Department of Psychiatry, University of Helsinki, Finland. His

research interests include developmental neuropsychiatric disorders and sleep disorders in psychiatry.

Laura Hokkanen, PhD, is an adjunct professor of clinical neuropsychology in the Institute of Behavioural Sciences, Division of Cognitive and Neuropsychology at the University of Helsinki, Finland. Her research interests include learning disabilities, memory disorders, and dementia.