

French Version of the Strengths and Weaknesses of ADHD Symptoms and Normal Behaviors (SWAN-F) Questionnaire

Philippe Robaey MD, PhD^{1,2}; Devendra Amre MBBS, PhD^{1,4}; Russell Schachar MD⁵;
Louise Simard PhD^{1,4}

Abstract

Objective: To evaluate internal and external consistency of a French adaptation of the SWAN (a 7-point rating strength-based scale, from far below to far above average) and its accuracy as a diagnostic test among children with Attention Deficit/Hyperactivity Disorder (ADHD). **Method:** Parents of 88 children referred for ADHD were interviewed using the SWAN-F, a structured interview (DISC-4.0) and the Conners' Rating Scale. Internal consistency and divergent and convergent validity of the SWAN-F were examined using the DISC-4.0 and Conners' Rating Scales as reference standards for four dimensions: Inattention, Hyperactivity/Impulsivity, ADHD, Oppositional Defiant Disorder. **Results:** The internal consistency of SWAN-F was within acceptable ranges for all dimensions (Cronbach's alpha greater than 0.80). Scores of the SWAN-F subscales were strongly associated with the DISC-4.0 diagnostic assignments and Conners' Rating Scales, following logical patterns of correspondence between diagnoses. Its accuracy as a diagnostic test was comparable to Conners' Rating Scale, with a lower rate of false positives. **Conclusions:** The information gathered with the SWAN-F is compatible with that obtained using the DISC-4.0 and Conners' Rating Scale. Strength-based rating scales have the potential to evaluate the normal distribution of behaviors and to provide reliable cut-off defining abnormal behavior.

Key words: Pharmacogenetic, Attention Deficit/Hyperactivity Disorder, Methylphenidate Treatment

Résumé

Introduction: Évaluer la cohérence du questionnaire SWAN-F lorsqu'il est administré à des enfants atteints de trouble du déficit d'attention avec hyperactivité (TDAH). **Méthodologie:** Les parents de 88 enfants qui ont reçu un diagnostic de TDAH ont rempli le SWAN-F, le DISC-4.0 et l'échelle d'évaluation Conners. Les sous-échelles du SWAN-F étaient basées sur une cotation en sept points qui classait les symptômes du DSM-IV de « nettement inférieurs à la moyenne » à « nettement supérieurs à la moyenne » chez des enfants d'âge identique. La cohérence, la validité discriminante et la validité convergente du SWAN-F ont été analysées au moyen du DISC4.0 et des échelles d'évaluation Conners qui ont servi de référence pour quatre dimensions: inattention, hyperactivité/impulsivité, TDAH, trouble oppositionnel. **Résultats:** Pour toutes les composantes du SWAN-F, la consistance interne se situaient dans la zone acceptable (le coefficient alpha de Cronbach était supérieur à 0,80). Les scores des sous-échelles du SWAN-F étaient fortement associés avec les diagnostics du DISC4.0 et aux scores de l'échelle d'évaluation Conners, correspondant logiquement aux diagnostics. **Conclusions:** Les données recueillies au moyen du SWAN-F sont compatibles avec celles obtenues au moyen du DISC-4.0 et de l'échelle d'évaluation Conners. Les échelles basées sur l'évaluation des forces pourraient être utilisées pour l'évaluation quantitative des symptômes dans les études longitudinales et génétiques et pour mesurer la réponse au traitement.

Mots clés: pharmacogénétique; trouble du déficit d'attention avec hyperactivité; traitement au méthylphénidate

Introduction

Most clinical rating scales quantify behaviors on a Likert scale anchored by standard descriptors, e.g., going from “never” to “very often”, through “sometimes” and “often”. When a weakness frequency (or intensity) is rated beyond “normal” limits, it becomes a symptom. Symptoms can thus be observed in subjects considered as healthy, as long as their number is beyond threshold and as long as they do not cause a significant impairment. For example, each ADHD behavioral descriptor has to be inappropriate for the child developmental level to be considered as a symptom. The diagnosis requires a symptom count above threshold and the associated impairment must be clinically significant, not specific to a situation (American Psychiatric Association, 1994). Normal variability is thus crucial at each level of the diagnosis process. However, in a pathological perspective,

normality is only defined by an absence or a low level of symptoms. This perspective creates problems as the distributions of symptoms are highly skewed and truncated in the normal population and as statistical cutoffs are generally based on the assumption of a normal distribution. For example, in an epidemiological sample, nearly 80% of children had scores of 1 or 0 (Just a Little or Not at All) for the ADHD items of the SNAP-IV. As a consequence, small changes in

¹Research Centre of Sainte-Justine Hospital, Montreal, Quebec

²Department of Psychiatry, Children's Hospital of Eastern Ontario, Ottawa, Ontario

⁴Department of Pediatrics, University of Montreal, Montreal, Quebec

⁵Department of Psychiatry, Hospital for Sick Children, Toronto, Ontario

Corresponding email: probaey@cheo.on.ca

Submitted: February 21, 2007; Accepted: May 3, 2007

cutoffs can have quite dramatic effect in the number of children above threshold (Swanson et al, 2005). Conversely, patients who show a low but near-threshold level of symptoms are defined as in remission. As it is not possible to determine where these children are placed within the normal distribution before and after treatment, therapeutic effects could also be overestimated.

These problems led some researchers to propose full-range rating scales. For example, Waschbusch and Sparkes (2003) designed an ADHD rating scale based on the assumption that symptoms can be rated below or above "normal" level. For example: "Does not seem to listen when spoken to directly" is rated with five anchor points from "much less" to "much more" than other (normal) children of the same age and sex. This attempt to break down the "Just a Little" and "Not at All" rating into categories from "much less" to "less" to "same" yielded a more normal distribution, although still negatively skewed, the "much less" rating being significantly overrepresented, especially in teacher ratings. Also, the proportion of children rated as "more" or "much more" symptomatic was strongly decreased (Waschbusch & Sparkes, 2003). Such a symptom-based full-range scale has the inherent difficulty of evaluating the relative rate of occurrence of behavioral descriptors that are by definition infrequent. A "much less" rating is also quite broad as it encompasses different levels of the corresponding strength. On the other side, evaluating the opposite strengths is likely to be easier and more reliable, as parents and teachers observed them much more frequently and generally define a weakness as a lack of strength.

The SWAN (Strengths and Weaknesses of ADHD-symptoms and Normal-behaviors ratings Scale - SWAN) is a scale where the informant is asked to assess the strength opposite to the ADHD symptoms (Swanson et al, 2005). Without changing the content, each item of the SNAP-IV rating scale (Swanson et al, 2001) was re-worded in order to capture the strength corresponding to the weakness. For example: "Often does not seem to listen when spoken to directly" becomes "Listen when spoken to directly" and is rated from "far below average" (-3) to "far above average" (+3) relative to chil-

dren of the same age. This approach yielded a normal distribution of ADHD scores (Hay et al, 2007) with a small positive skewness, the above average rating being somewhat overrepresented. The proportion of children rated as symptomatic (below and far below average) was also decreased.

The objectives of this study was to test internal and external consistency of a French adaptation of the SWAN and its accuracy as a diagnostic test, by using a symptom-based interview (DISC-4.0) and different problem-based scales (Conner's Parent and Teacher Rating Scale) in a sample of children referred for ADHD in a university-based specialized clinic in Montreal, Canada.

Methods

Participants

Children of 6-9 years of age (n=124) referred by their physicians with a suspected diagnosis of ADHD were recruited through the Interdisciplinary Research Program on Hyperactivity at Sainte-Justine Hospital (Montreal, Canada). Children with an IQ of less than 80 (Wechsler, 1991), born prematurely (<35 weeks of gestation), with severe learning or language disabilities and with neurological diseases (e.g. epilepsy) were excluded. Following Ethics Board approval, informed consent was obtained from all participating families.

Measures

DISC-4.0: During the scheduled hospital visit, the parent was administered a structured computerized interview, the Diagnostic Interview Schedule for Children 4.0 (DISC-4.0), by a trained interviewer. The DISC-4.0 has been translated into French by two independent research teams in Montreal. These French versions have been compared systematically and standardized to build the DISC-4.0 version used in this study. The French DISC-4.0 has not been specifically validated, but the test-retest reliabilities of the English version of the DISC-4.0 (Shaffer et al, 2000) and of the previous French version (DISC2.3) (Breton et al, 1998) are considered satisfactory. Diagnoses of Attention Deficit Hyperactivity Disorder (ADHD), Conduct Disorder (CD), and Oppositional Defiant Disorder (ODD) were obtained from the

DISC-4.0 computerized algorithms. The presence (or not) of each ADHD symptom and the number of positive symptoms rated with the DISC-4.0 were evaluated.

SWAN-F: The original SWAN is based on the DSM-IV criteria and is available through <http://www.adhd.net/>. In the SWAN, symptoms from the DSM-IV criteria list were reworded using a strength-based formulation; for example, “often has difficulty sustaining attention in tasks or play activities” was modified to “is able to sustain attention in tasks or play activities”. We adapted the SWAN, as the original version retained some symptom-based explanations and translated it in French (using a back-translation): see appendix. The SWAN-F includes the items for the DSM symptoms for Inattention (9 items), Hyperactivity/ Impulsivity (9 items), Oppositional Defiant Disorder (8 items) but not the 3 additional “sluggish tempo” items and a 30th item “Avoid quarrelling” of the downloadable SWAN. Three Conduct Disorder items and 5 prosocial items were also included, but not analyzed.

The SWAN-F items are scored according to a seven-point scale ranging from “far below average” (-3) to “far above average” (+3) relative to children of the same age, a score of 0 is “in the average”. For each child, scores for SWAN-F Inattention (S-IN scores) and Hyperactivity/Impulsivity (S-HY/IM scores) subscales were calculated as the average of the ratings obtained for the 9 inattention and the 9 hyperactivity/impulsivity items, respectively. The overall SWAN-F ADHD score (S-ADHD scores) was the average of the 18 ratings. Scores for the SWAN-F Oppositional Defiant Disorder (S-ODD scores) was estimated as the average of the ratings obtained for the 8 ODD items. Two to three weeks prior to the scheduled evaluation at Sainte-Justine Hospital, the parent (84% mothers) received and completed the SWAN-F.

Revised Conners’ Parents and Teachers Rating Scales (CPRS-R and CTRS-R; Conners et al, 1998a,b): The French version of the Conners’ Parent and Teacher Revised Rating Scales (long version) were administered to the parent and the teacher of the child respectively. As age or sex-corrected scores are not available for the SWAN or the SWAN-F, and as the age range was small, the raw ratings of the SWAN-F scales were compared with the raw

scores of the Conners’ subscales, rather than the usual standardized T-scores. Conners’ subscales were calculated as weighted addition (total of the ratings x number of items in the scale/number of completed items) of the respective items rated from zero to three.

General Information Questionnaire: Information on epidemiological, socio-demographic and medical variables was collected using a structured questionnaire addressed to the parent. Specifically, information on the age of the child, sex, mother’s and father’s education, Canadian origin and family income and structure was collected.

Wechsler Intelligence Scale for Children third edition (WISC-III) (Wechsler, 1991): This scale was used by the psychologist to assess the IQ of the child to determine if the child met inclusion/exclusion criteria for this study.

Data Analyses

Preliminary analyses involved studying the association between SWAN-F scores and socio-demographic variables such as age (in months), sex, mother’s and father’s education (less than college/college or university level), Canadian origin (one or two parents born outside Canada/both parent born in Canada), family income (<35 000\$CAN/≥35 000\$CAN) or family structure (both biological parents vs. other: mono-parental/reconstituted). Internal consistency of the SWAN-F subscales was assessed by estimating the Cronbach’s alpha and correlations coefficients of each item with the corresponding scale (item-total correlation). To investigate external consistency, student t-tests were performed to compare mean SWAN-F scores between children with and without ADHD, ODD and CD, diagnosed using the DISC-4.0. SWAN-F scores within specific ADHD subtypes (DISC-4.0) were compared using one-way ANOVA. Pearson correlation coefficients were calculated to assess the association between Conners’ and SWAN-F scores. Finally, using DISC-4.0 ADHD diagnostic assignment as reference, Receiver Operating Characteristic curves (or ROC curves) (Beck and Shultz, 1986) were generated to investigate the capacity of the Conners’ Parent Rating Scale and SWAN-F ADHD scores to discriminate ADHD cases versus non-cases. SPSS 15.0 was used for the analyses.

Table 1: Internal consistency of SWAN-F subscales (N=88)

Subscales	Number of items	Cronbach's Alpha coefficients	Range of item-total correlation coefficients
S-IN	9	0.89	0.31-0.77
S-HY/IM	9	0.88	0.33-0.80
S-ADHD	18	0.91	0.28-0.71
S-ODD	8	0.88	0.56-0.77

Table 2: Comparison of mean SWAN-F scores between children with and without ADHD, ODD and CD

DISC-4.0 diagnosis	SWAN-F subscales			
	Mean Scores (Standard Deviation)			
	S-IN	S-HY/IM	S-ADHD	S-ODD
Non-ADHD (n=14)	-0.61 (0.72)	0.01 (0.90)	-0.30 (0.67)	-0.35 (0.97)
ADHD (n=74)	-1.59 (0.84)***	-1.16 (0.90)***	-1.38 (0.74)***	-0.85 (1.04)
Non ODD (n=43)	-1.27 (0.74)	-0.56 (0.90)	-0.91 (0.72)	-0.12 (0.89)
ODD (n=45)	-1.59 (1.00)	-1.37 (0.93)***	-1.48 (0.83)**	-1.38 (0.75)***
Non-CD (n=63)	-1.34 (0.89)	-0.77 (1.00)	-1.06 (0.80)	-0.55 (1.00)
CD (n=25)	-1.66 (0.88)	-1.48 (0.82)**	-1.57 (0.78)**	-1.31 (0.93)**

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: DISC-4.0 diagnosis: ADHD; Attention Deficit Hyperactivity Disorder, ODD; Oppositional Defiant Disorder, CD; Conduct Disorder. SWAN-F scales: S-IN; Inattention, S-HY/IM; Hyperactivity/Impulsivity, S-ADHD; Attention Deficit Hyperactivity Disorder, S-ODD, Oppositional Defiant Disorder.

Results

Among the 124 referred families, 94 (76%) accepted to participate and 6 children were excluded because they presented with an IQ of less than 80. Thus, the study sample consisted of 88 subjects; 68 (78%) boys and 20 (22%) girls. The age distribution of the children was as follow: 19%, 32%, 34% and 15% were 6, 7, 8 and 9 years of age, respectively. Both parents were of Canadian origin for 76% of the children. College education was completed by 61% of the fathers and 56% of the mothers.

Among the 88 children investigated, 74 (84%) were diagnosed with ADHD according to DISC-4.0. Of these, 26 (35%) were categorized as Inattentive, 15 (20%) as Hyperactive/Impulsive and 33 (45%) as Combined type. A total of 57 (77%) cases presented at least one other psychiatric disorder; 46 (62%) presented with CD or ODD, 24 (32%) with at least one Mood or Anxiety Disorder, 10 (14%) with Tic or Tourette's Disorder and 4 (5%) with Elimination or Eating Disorders. Forty nine subjects (56%) were currently being treated with psychostimulants, this frequency was similar among subjects classified as ADHD and

non-ADHD according to DISC-4.0.

The SWAN-F scores were normally distributed within subscales among the study population. Multiple linear regression revealed no association between any of the socio-demographic variables and the different SWAN-F scores. However, the modes of the S-ADHD scores distribution, were very similar (Mean=-1.22; S.D. = .77 in boys and Mean=-1.15; S.D.= .85 in girls), as the medians (-1.36 for boys and -1.0 for girls), the modes of the S-ADHD scores distribution were -2.06 in boys and only -0.78 in girls.

Internal consistency

Internal consistency of SWAN-F was within an acceptable range for all subscales (Table 1), yielding coefficients above 0.80 (Nunnally & Bernstein, 1994). A Cronbach's alpha coefficient of 0.91 was observed for the S-ADHD subscale (18 ADHD items). While all other item-correlations with the S-IN subscale were above 0.53, the item "listen when spoken directly" (DSM-IV 1c ADHD item; American Psychiatric Association, 1994) had an item-total correlation of 0.31. The "able to talk with a normal flow" item (DSM-IV 2f ADHD item; American

Psychiatric Association, 1994) had a coefficient of 0.33 when item-correlation with S-HY/IM subscale was investigated. Coefficients were greater than 0.51 for the remaining items. The item-total correlations for the S-ODD subscale were all above 0.64.

External consistency

Table 2 presents the mean SWAN-F scores for children with and without ADHD, ODD and CD, respectively. Significant differences in mean SWAN-F scores between ADHD and non-ADHD cases were observed for S-IN ($t=4.08$, $p<0.001$), S-HY/IM ($t=4.44$, $p<0.001$) and S-ADHD ($t=5.06$, $p<0.001$) subscales, with children with ADHD presenting lower scores (more impaired) than non-ADHD children. When ODD and non-ODD children were compared, significant differences between mean SWAN-F scores were observed for all scales except the S-IN subscale, with children with ODD presenting lower scores (more impaired) than non-ODD children. Finally, significant differences in mean scores of the S-HY/IM, S-ADHD and S-ODD were observed between children with and without CD. Children with ODD and CD presented lower SWAN-F scores than non-impaired children.

ANOVA results evaluating the consistency between SWAN-F scores and ADHD subtypes are presented in Table 3. Predictably, mean scores of the S-IN subscale were significantly

lower for children with Inattentive or Combined subtypes compared to non-ADHD children and children presenting with the Hyperactive/Impulsive subtype. Consistently, mean scores of the S-HY/IM subscale were significantly lower for children with Hyperactive/Impulsive or Combined subtypes compared to non-ADHD and Inattentive children. These specific patterns were no longer observed for the mean global S-ADHD scores. S-ADHD scores were significantly lower for children with the combined subtype compared to children with each specific subtype and for children with ADHD, whatever the subtype, compared to non-ADHD children.

The mean SWAN-F ratings for each specific DSM-IV item were compared between children with or without the respective symptoms as assessed with the DISC-4.0. The mean SWAN-F ratings were significantly lower ($p<0.05$) among children with the symptoms for 7 of the 9 Inattentive (“listen when spoken to directly” $p=.15$, and “organize work and activities”; $p=.08$ in a bilateral test) and all the Hyperactive/Impulsivity symptoms and the S-HY/IM score ($r= -0.69$, $p<0.001$): the more negative the scores, the more numerous the symptoms.

When studying the relationship between ratings of the SWAN-F and Conners' Parent

Table 3: Comparison of SWAN-F scores between ADHD subtypes.

DISC-4.0 diagnosis	S-IN scores		S-HY/IM scores		S-ADHD scores	
	Mean (SD)	Group differences	Mean (SD)	Group differences	Mean (SD)	Group differences
Non-ADHD (n=14)	-0.61 (0.71)	Non-ADHD vs, IN***, Non ADHD vs, COM**	0.01 (0.90)	Non ADHD vs, HY/IM**, Non ADHD vs, COM***	-0.30 (0.67)	Non ADHD vs, IN**, Non-ADHD vs, HY/IM*, Non ADHD vs, COM***
ADHD subtypes (N=74)						
Inattentive (IN) (n=26)	-1.76 (0.64)	IN vs, HY/IM**	-0.54 (0.81)	IN vs. COM***	-1.15 (0.65)	IN vs, COM*
Hyperactive/Impulsive (HY/IM) (n=15)	-0.92 (1.18)	HY/IM vs. COM**	-1.17 (0.93)		-1.05 (0.92)	HY/IM vs, COM*
Combined (COM) (n=33)	-1.77 (0.65)		-1.63 (0.66)		-1.70 (0.59)	

* $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Note: SWAN-F scales: S-IN; Inattention, S-HY/IM; Hyperactivity/Impulsivity, S-ADHD; Attention Deficit Hyperactivity Disorder.

Table 4: Item-by-item Pearson correlation coefficients between SWAN-F scores and Conners' Parent Inattention or Hyperactivity/Impulsivity subscale scores.

Conners' Parent Inattention scores	Correlation between specific item SWAN-F scores and		
	Conners' Parent Inattention scores	Conners' Parent Hyperactivity/Impulsivity scores	
a) attending to detail	-0.44***	a) sitting still	-0.65***
b) sustaining attention	-0.51***	b) staying seated	-0.68**
c) listening	-0.14	c) modulating motor activity	-0.56***
d) following through	-0.52***	d) playing quietly	-0.58***
e) organizing	-0.60***	e) settling down	-0.71***
f) engaging in sustained effort	-0.61***	f) modulating verbal activity	-0.47***
g) keeping track of things	-0.59***	g) reflecting on questions	-0.54***
h) ignoring extraneous stimuli	-0.54***	h) awaiting turn	-0.59***
i) remembering	-0.47***	i) entering into others activities	-0.66***
Inattention subscale	-0.79***	Hyperactivity/impulsivity subscale	-0.85***

*p<0.05, ** p<0.01, ***p<0.001

NB: all correlations are negative as impairment is rated as more negative in SWAN-F scores and more positive in Conners' subscales

Rating Scale, we observed significant correlations for Conners' Parent subscale scores and all specific SWAN-F scores except the "listen when spoken directly" item (Table 4). Among the 45 teachers (51%) who completed the Conners-Teacher scale, significant correlations were observed between Conners' Teacher and Conners' Parent scores for the Inattention ($r=0.36$ $p<0.05$) and Hyperactivity/Impulsivity ($r=0.56$ $p<0.001$) subscales. Furthermore, significant correlations were also observed between SWAN-F and Conners' Teacher scores for the Inattention ($r= -0.30$ $p<0.05$) and Hyperactivity/Impulsivity ($r= -0.32$ $p<0.05$) subscales.

ROC curves were generated to evaluate the discriminant capacity of the Conners' Global Index and S-ADHD subscale, using the DISC-4.0 diagnosis assignment (gold standard) for any type of ADHD as reference. The estimated areas under the curves (AUC) were similar and significantly different from 0.5 (no possible discrimination) for the Conners' Global Index' (AUC= 0.79; CI 95%: 0.66-0.92) and SWAN-F scores (AUC= 0.89; CI 95%: 0.81-0.97). For the S-ADHD scores the optimal sensitivity (0.86) and specificity (0.88) corresponded to a cut-off of -0.60. For the Conners' Global Index, the optimal cut-off corresponded to a sensitivity of 0.75 and a specificity of 0.80. Increasing the sensitivity to 0.85 by lowering the cut-off resulted in a sudden drop in specificity (0.50),

an observation that suggests that the SWAN could yield a lower false positive rate than the Conners' Rating Scale near cut-off. Results were similar when using the DSM-IV global symptom scale of the Conners' Rating Scale: optimal sensitivity and specificity were 0.80, but any further increase in sensitivity generated a sudden decrease in specificity.

Discussion

Overall, the criteria used to assess the reliability and accuracy of the SWAN-F in this study sample showed that the SWAN-F was consistent with the DISC-4.0 and the CPRS-R. SWAN-F showed high internal consistency for all subscales. Meaningful patterns of correspondence were observed between S-IN, S-HY/IM and S-ADHD scores and ADHD, ODD and CD diagnoses, as well as with ADHD subtypes. Although high correlations were observed for almost all DSM-IV ADHD symptoms (item-by-item comparisons), the item "listen when spoken to directly" showed low internal as well as external consistency. The DSM-IV wording of the corresponding criterion is "often does not seem to listen when spoken directly". The lower consistency may result from a misunderstanding of the question by the parents who scored the symptom as one of "opposition" rather than "inattention". "Does not listen" ("N'écoute pas") is often used to mean "does not obey" in French. It may be appropriate to modify the SWAN-F symptom

question to “Fais attention quand on lui parle directement-Pay attention when spoken to directly”. One Hyperactivity/Impulsivity item worded “Quand il parle, est capable de régler son débit suivant le contexte” in the SWAN-F - for “modulate verbal activity (control excess talking)” in the SWAN (corresponding to “often talks excessively” as DSM-IV criterion) also showed low internal consistency. This may be related to the relative complexity of the item. A simpler wording of this item (e.g., “Se retient de trop parler, sur n’importe quoi- Keep themselves from talking too much, whatever the topic”) could possibly increase the observed internal and external consistencies.

The SWAN was developed in response to concerns that the SNAP-IV as well as the other available symptom-based truncated checklist may overestimate the number of youths with ADHD, because of the skewed distribution. The summary scores of the SWAN-F were normally distributed even in a clinical sample and thus departed from the J curve generated by one-tailed ADHD rating scales. The use of normally distributed ADHD ratings could also be of interest for other clinical and research issues in which the estimation of the normal variability is central.

For example, it may improve our understanding of the gender impact on ADHD expression. According to the polygenic multiple threshold model, girls are less likely than boys to be diagnosed with ADHD because girls require greater liability to manifest ADHD than boys. Mothers also perceived the DSM-IV ADHD, ODD, and CD criteria as more descriptive of boys (Ohan & Johnston, 2006). However, in order to determine sex-specific cut-off, the full sex-specific distribution should be known. In an epidemiological sample of 872 boys and 812 girls, more girls than boys seemed to have moderate symptomatic level using the SWAN (Manly et al, 2005). By using a symptomatic rating scale going from much less to much more than other children, Waschbusch & King (2006) found in an epidemiological sample of 781 boys and 710 girls that a small percentage of girls with a higher than average ADHD and ODD symptom count did not meet DSM-IV diagnostic threshold. In the present clinical sample, we found no significant difference according to gender for the mean SWAN-F scores. However,

the large difference in mode, more negative for boys than for girls, suggests an over-representation of near-threshold ratings in girls, even in a referred sample. Further research is thus needed to explore the full distribution of scores according to gender, and to develop age and gender norms for the SWAN.

The use of SWAN-F is also pertinent in genetic studies. In a twin design, the structural equation modeling is very sensitive to departure from normal distribution and truncated measures are by definition skewed. In addition, scoring individuals struggling with difficulties as well as those performing well above average increases the validity of the correlations within twin pairs. The direct comparison of truncated (Australian Twin Behaviour Rating Scale - ATBRS) with full-range scores (SWAN) suggests (Hay et al, 2007) that the proportion of children rated as having problems is inflated by using a problem-based truncated scale, as compared to a full-range strength-based scale. As the “not at all” descriptor is also used to describe a much wider range of behaviors than the other descriptors (normal but also different degrees of strength), parents seem to re-distribute their scores and to use more frequently the “negative” standard descriptors (from sometimes to very often). Not only is the highly skewed distribution problematic in genetic studies, but increasingly, association studies are using discordant or concordant pairs to detect linkages. Full-range questionnaires are obviously more appropriate to identify extremely discordant pairs than truncated ones and may be preferred (Cornish et al, 2005). The same reasoning holds for studies that look for an association between any biological or neuropsychological measures and behaviors by selecting subjects at both ends of the distribution of behavioral descriptors. For example, the SWAN was used to select children at the extremes of a “normal” ADHD continuum in a study on the relationship between rightward visuo-spatial bias and poor attention within the normal child population using the Line Bisection test (Manly et al, 2005).

Limitations

Some limitations of the study should be considered prior to interpretation. The accuracy of the SWAN-F was evaluated among ADHD

SWAN-F

	Beaucoup en dessous de la moyenne	En dessous de la moyenne	Un peu en dessous de la moyenne	Dans la moyenne	Un peu au dessus de la moyenne	Au dessus de la moyenne	Beaucoup au dessus de la moyenne	code
1 Est capable de se mettre d'accord avec les adultes.	-3	-2	-1	0	1	2	3	2
2 Est capable de s'amuser calmement (dans les jeux ou les activités de loisirs).	-3	-2	-1	0	1	2	3	1
3 Console un enfant qui pleure ou qui est bouleversé.	-3	-2	-1	0	1	2	3	3
4 Est capable de s'organiser dans ses travaux ou ses activités.	-3	-2	-1	0	1	2	3	0
5 Est capable de s'arrêter et de rester tranquille.	-3	-2	-1	0	1	2	3	1
6 Est capable de garder son sang froid, de rester calme.	-3	-2	-1	0	1	2	3	2
7 Selon le contexte, est capable de rester assis sans se lever (en classe ou ailleurs).	-3	-2	-1	0	1	2	3	1
8 Évite de bousculer, de menacer ou d'intimider les autres.	-3	-2	-1	0	1	2	3	4
9 Est capable de se joindre sans s'imposer à une conversation ou à un jeu.	-3	-2	-1	0	1	2	3	1
10 Reconnaît la responsabilité de ses erreurs et de ses mauvaises conduites.	-3	-2	-1	0	1	2	3	2
11 Évite d'être impliqué dans des bagarres.	-3	-2	-1	0	1	2	3	4
12 Écoute quand on lui parle personnellement.	-3	-2	-1	0	1	2	3	0
13 Aide un enfant qui se sent malade, qui s'est blessé.	-3	-2	-1	0	1	2	3	3
14 Est capable de faire attention aux détails et d'éviter de faire des fautes d'étourderie.	-3	-2	-1	0	1	2	3	0
15 Est capable de suivre les consignes et de terminer ses devoirs/tâches.	-3	-2	-1	0	1	2	3	0
16 Est capable d'oublier sa rancune et de ne pas vouloir se venger.	-3	-2	-1	0	1	2	3	2
17 Est capable de veiller sur les choses nécessaires à son travail (livres, crayons, etc.) ou à ses activités (jouets, etc.).	-3	-2	-1	0	1	2	3	0
18 Ne se laisse pas froisser, vexer ou ennuyer par les autres.	-3	-2	-1	0	1	2	3	2
19 Ne se sent pas victime d'injustice et contrôle sa colère.	-3	-2	-1	0	1	2	3	2
20 Est capable de soutenir son attention au travail ou dans les jeux.	-3	-2	-1	0	1	2	3	0
21 Est capable d'entreprendre des tâches qui nécessitent un effort mental soutenu (à l'école ou à la maison).	-3	-2	-1	0	1	2	3	0
22 Dans la vie quotidienne, est capable de retenir ce qu'on lui dit ou ce qu'il doit faire.	-3	-2	-1	0	1	2	3	0
23 Invite un enfant qui se tient à l'écart à se joindre à son groupe de jeu.	-3	-2	-1	0	1	2	3	3
24 Selon le contexte, est capable de se retenir de courir ou de grimper, (n'a pas "des fourmis dans les jambes").	-3 1	-2	-1	0	1	2	3	3
25 Évite de faire délibérément des choses qui pourraient fâcher les autres.	-3	-2	-1	0	1	2	3	2
26 Est capable de se concentrer, de ne pas se laisser distraire par des stimuli externes.	-3	-2	-1	0	1	2	3	0
27 Évite de commencer les bagarres.	-3	-2	-1	0	1	2	3	4
28 Attends qu'une question soit entièrement posée avant d'y répondre.	-3	-2	-1	0	1	2	3	1
29 Accepte de suivre les règles et de répondre aux demandes des adultes.	-3	-2	-1	0	1	2	3	2
30 Devant une querelle ou une dispute, essaie de l'arrêter	-3	-2	-1	0	1	2	3	3
31 Est capable de rester assis sans s'agiter, en contrôlant les mouvements de ses mains et ses pieds.	-3	-2	-1	0	1	2	3	1
32 Quand il parle, est capable de régler son débit suivant le contexte.	-3	-2	-1	0	1	2	3	1
33 Est capable d'attendre son tour	-3	-2	-1	0	1	2	3	1
34 Se propose pour aider à nettoyer un dégât fait par quelqu'un d'autre.	-3	-2	-1	0	1	2	3	3

Scoring: S-IN: average item code 0; S-HY: average items code 1; S-ODD: average code 2; S-PRO: average code 3; S-CD: average code 4.

patients referred to clinicians as part of an ongoing research program. Thus the population was in many ways a “selected population” and extrapolation of findings to the “general” or similar “ADHD” populations may be inappropriate. Further studies among larger unselected populations will be necessary to further evaluate the utility of the SWAN-F. Although, information gathered with the SWAN-F significantly correlated with that obtained by the Conners’ Teachers Rating Scale, the low response rate among the teachers may have influenced the results. In the present study, classification as ADHD or non-ADHD was based on the findings obtained using DISC-4.0. Diagnosis using DISC-4.0 has inherent limitations. Information is collected from only 1 informant and judging the exactness of the information is not possible.

Clinical implications

To our knowledge, this is the first validation study of the SWAN assessing internal and external consistency in a referred sample. Although, the investigation of their psychometric properties needs to be further pursued, the SWAN and SWAN-F could nonetheless in their present form have good potential for use in clinical and research setting because they retain the advantages of other rating scales: simple to comprehend, rapidly completed, providing quantitative scores. In addition, as they are based on strengths, they allow parents to recognize them when they exist and thus may decrease guilt and stigmatization associated with reporting the child’s difficulties. Moreover, these scales may allow limiting some bias in clinical decision-making. A study by Lewczyk et al (2003) showed that the poor concordance among clinicians when diagnosing ADHD and disruptive disorder partly resulted from the concern to avoid false positives, even at the cost of increasing the risk for false negatives. The Conner’s Parent Rating Scale was reported to have a sensitivity of 0.92 and a specificity of 0.94 (Conners et al, 1998a), in separating children with a confirmed diagnosis of ADHD and an epidemiological sample (mean age = 10.16 years; SD = 3.40). However, the present data suggest that the accuracy could be lower in a clinically referred sample (around 0.80), comparable to the SWAN. Moreover, if further studies confirm the lower false positive rate

associated with the SWAN-F, it could constitute a more reliable tool for clinicians than symptom-based scales.

Acknowledgements

Study statistical expert: Devendra Amre has acquired doctoral training in Epidemiology and Biostatistics. This study was funded by the “Réseau de santé mentale du Fond de la Recherche en Santé du Québec” and Sainte-Justine Hospital Research Centre. We would like to express our gratitude to the children and their families who participated in the study and to François L’Heureux for his daily support in coordinating this study.

References

- American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (DSM-IV). Washington, DC: American Psychiatric Association.
- Beck J.R. & Shultz E.K. (1986). The use of relative operating characteristic (ROC) curves in test performance evaluation. *Archives of Pathology & Laboratory Medicine*, 110:13-20.
- Breton J.J., Bergeron L., Valla J.P., Berthiaume C., St Georges M. (1998). Diagnostic Interview Schedule for Children (DISC-2.25) in Quebec: reliability findings in light of the MECA study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 37:1167-1174.
- Conners C.K., Sitarenios G., Parker J.D., Epstein J.N. (1998a). The revised Conners' Parent Rating Scale (CPRS-R): factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology*, 26:257-68.
- Conners C.K., Sitarenios G., Parker J.D., Epstein J.N. (1998b). Revision and restandardization of the Conners' Teacher Rating Scale (CTRS-R): factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology*, 26:279-91.
- Cornish K.M., Manly T., Savage R., Swanson J., Morisano D., Butler N., Grant C., Cross G., Bentley L., Hollis C.P. (2005). Association of the dopamine transporter (DAT1) 10/10-repeat genotype with ADHD symptoms and response inhibition in a general population sample. *Molecular Psychiatry*, 10(7):686-98.
- Hay D.A., Bennett K.S., Levy F., Sergeant J., Swanson J. (2007). A twin study of attention-deficit/hyperactivity disorder dimensions rated by the strengths and weaknesses of ADHD-symptoms and normal-behavior (SWAN) scale. *Biological Psychiatry*, 61(5):700-5.
- Lewczyk C.M., Garland A.F., Hurlburt M.S., Gearing J., Hough R.L. (2003). Comparing DISC-IV and clinician diagnoses among youths receiving public mental health services. *Journal of the American Academy of Child and Adolescent Psychiatry*, 42(3):349-56.
- Manly T., Cornish K., Grant C., Dobler V., Hollis C. (2005). Examining the relationship between rightward visuo-spatial bias and poor attention within the normal child population using a brief screening task. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 46(12):1337-44.
- Nunnally J.C., Bernstein I.H. (1994). *Psychometric Theory*, 3rd edition. New York: McGraw-Hill.

- Ohan J.L. & Johnston C. (2006). Gender appropriateness of symptom criteria for attention-deficit/hyperactivity disorder, oppositional-defiant disorder, and conduct disorder. *Child Psychiatry and Human Development*, 35(4):359-81.
- Shaffer D., Fisher P., Lucas C.P., Dulcan M.K., Schwab-Stone M.E. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC- IV): description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39: 28-38.
- Swanson J.M., Kraemer H.C., Hinshaw S.P., Arnold L.E., Conners C.K., Abikoff H.B., Clevenger W., Davies M., Elliott G.R., Greenhill L.L., Hechtman L., Hoza B., Jensen P.S., March J.S., Newcorn J.H., Owens E.B., Pelham W.E., Schiller E., Severe J.B., Simpson S., Vitiello B., Wells K., Wigal T., Wu M. (2001). Clinical relevance of the primary findings of the MTA: success rates based on severity of ADHD and ODD symptoms at the end of treatment. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40:168-79.
- Swanson J., Schuck S., Mann M., Carlson C., Hartman K., Sergeant J., et al (2005): Categorical and dimensional definitions and evaluations of symptoms of ADHD: The SNAP and the SWAN Ratings Scales [Draft]. Available at: http://www.adhd.net/SNAP_SWAN.pdf.
- Waschbusch D. and Sparkes S. (2003). Rating scale assessment of Attention Deficit/Hyperactivity Disorder (ADHD) and Oppositional Defiant Disorder (ODD): Is there a normal distribution and does it matter? *Journal of Psychoeducational Assessment*, 21, 261-281.
- Waschbusch D. & King S. (2006). Should sex-specific norms be used to assess Attention-Deficit/Hyperactivity Disorder or Oppositional Defiant Disorder? *Journal of Consulting and Clinical Psychology*, 74(1),179-185.
- Wechsler, D. (1991). Manual for the Wechsler Intelligence Scale for Children (3rd ed.). New York: Psychological Corporation.