Published in final edited form as:

J Atten Disord. 2013 February ; 17(2): 152–162. doi:10.1177/1087054711427399.

The SWAN Captures Variance at Both the Negative and Positive Ends of the ADHD Symptom Dimension

Anne B. Arnett,
University of Denver

Bruce F. Pennington,
University of Denver

Angela Friend,
University of Colorado, Boulder

Erik Willcutt,
University of Colorado, Boulder

Brian Byrne,
University of New England, Australia, and the University of Linköping, Sweden

Stefan Samuelsson, and
University of Linköping, Sweden, and Stavanger University, Norway

Richard K. Olson
University of Colorado, Boulder and the University of Linköping, Sweden

Abstract

Objective—The Strengths and Weakness of ADHD and Normal Behavior Scale (SWAN) differs from previous parent-reports of Attention Deficit Hyperactivity Disorder (ADHD) in that it was designed to also measure variability at the positive end of the symptom spectrum. The SWAN was compared to another ADHD measure that focuses on deficits associated with ADHD.

Method—The psychometric properties of the SWAN were tested and compared to an established measure of ADHD, the Disruptive Behavior Rating Scale (DBRS) in samples of school age twins in the US and Australia.

Results—The SWAN demonstrates comparable validity, reliability and heritability to the DBRS. Plots of the scales reveal heteroscedasticity supporting the SWAN as a preferred measure of both negative and positive ADHD behaviors.

Conclusion—The additional variance captured by the SWAN at the adaptive ends of the ADHD symptom dimensions makes it a promising tool for behavioral and molecular genetic studies of ADHD.

The Strengths and Weaknesses of ADHD Symptoms and Normal Behavior (SWAN, Swanson et al., 2006) questionnaire is a relatively new parent-report measure of symptoms associated with Attention Deficit/Hyperactivity Disorder (ADHD). Unlike other established ADHD questionnaires, such as the Disruptive Behavior Rating Scale (DBRS, Barkley & Murphy, 1998), the SWAN was constructed with the intent of capturing variability at the good ends of attention and activity regulation as well as at the symptomatic levels of these dimensions. For example, the SWAN asks how well the child “listens when spoken to” (attention) or “awaits turn” (activity regulation).

The overall goals of this study were to evaluate the psychometric properties of the SWAN and compare it to the DBRS, since the SWAN could be useful in research aimed at

understanding strengths as well as weaknesses in attention and impulse control. Contemporary behavioral-genetic and phenotypic studies of ADHD conceptualize the diagnosis as an extreme phenotype that is not etiologically distinct from the spectrum of symptoms that fall within the normal range (Pennington, 2002). On the other hand, it is possible that the spectrum of symptoms that are behaviorally adaptive do not share a common etiology with the disorder. Measuring variance in both the positive as well as the negative ends of the symptom dimensions can address this question.

Background

Attention Deficit/Hyperactivity Disorder (ADHD) is one of the most common neurodevelopmental disorders, affecting approximately 4% of 4–8 year olds in the general population (Center for Disease Control, 2005). Current diagnostic guidelines include two correlated symptom dimensions of Inattention (I) and Hyperactivity/Impulsivity (HI). A child with ADHD may be elevated on one or both symptom dimensions, resulting in three subtypes of ADHD: Inattentive, Hyperactive-Impulsive, and Combined. In order to meet diagnostic criteria for ADHD, symptoms must be present before age seven and occur in more than two settings, be inconsistent with developmental stages, and cause impairments in adaptive functioning (American Psychiatric Association, 1994; 2000). Parent- and teacher-report measures, such as the DBRS, typically show internal and external validity for symptoms of inattention and hyperactivity/impulsivity (e.g., Bussing et al., 2008; Friedman-Weieneth, Doctoroff, Harvey, & Goldstein, 2009). However, these same scales do not capture variation at the positive end of the two symptom dimensions (Polderman et al., 2007). Studies of genetic etiologies for ADHD have found that heritability for these symptoms is consistent across the continuum of severity (Stevenson et al., 2005). However, without a scale that measures variance at the adaptive end of the phenotype, it is difficult to estimate the degree to which these genetic factors are shared at both ends of the symptom dimensions. A full spectrum of genetic variance in these symptoms will allow for more precise evaluation of the specific genetic and neuropsychological correlates of attention and impulse regulation.

The “Strengths and Weaknesses of ADHD Symptoms and Normal Behavior Scale” (SWAN, Swanson et al., 2006) was developed with the goal of detecting variance at both the negative and adaptive ends of the two symptom dimensions. The SWAN differs from other ADHD measures in its balanced Likert scale that includes positively-phrased, DSM-IV anchored items (see instrument descriptions below). However, despite this apparent face validity, few studies have examined its psychometric properties, particularly whether the SWAN truly captures meaningful variance at the favorable end of the Inattention/HI distribution. Polderman, et al. (2007) plotted SWAN scores with scores on the Attention Problems scale of the Child Behavior Checklist (CBCL, Achenbach, 1991) to test for heteroscedasticity, or a difference in variances between the two scales. While the SWAN showed variance at the positive end of Inattention and HI levels, CBCL “positive” scores were truncated at zero. However, this study did not test the external validity of the SWAN scores. Moreover, unlike the SWAN, the CBCL ADHD scale comprises 11 items that do not map directly onto DSM-IV ADHD criteria, but instead tap global symptoms of inattention and hyperactivity-impulsivity. For example, item 10 on the CBCL reads, “Can’t sit still, restless, or hyperactive,” whereas the DSM-IV criteria and the SWAN identify nine specific behaviors that fall under the umbrella term, “hyperactive.” So, for both these reasons the Polderman et al. (2007) results are not conclusive.

Young, Levy, Martin and Hay (2009) found that the SWAN identified a distribution of ADHD subtypes consistent with the Rasch model, suggesting consistency in the relative predictive strength of individual items on the measure; however, the authors did not correct

for inflation due to familial correlations. Finally, Hay, et al. (2007) compared the SWAN to ADHD items on the DSM-IV oriented Australian Twin Behavior Rating Scale in a group of Australian twins, age 6–17 years. This study confirmed that the SWAN provides a more normal distribution as compared to the skew and kurtosis found in a 0–3 rating scale similar to the DBRS, but other validity tests were not conducted.

In studies of genetic and environmental influences on inattention and impulse control, the ability to measure individual differences across the whole distribution of these symptom dimensions would be useful. Thus, the goal of the current study is to examine the psychometric properties of the SWAN to determine whether it can be useful in future research. Specifically, the goals of the study were to examine:

1. Internal validity of the SWAN as measured by internal consistency and test-retest reliability, exploratory factor analyses, and tests of normality.
2. Convergent validity of the SWAN with the DBRS at the symptom severity level, via sensitivity/specificity to identify clinically elevated children, and by estimating heritability for symptoms of ADHD.
3. External validity of the SWAN, as compared to the DBRS, as measured by phenotypic and genetic correlations with a cognitive latent factor and ratings of interference with daily functioning.

Method

Participants

Participants in this study were 379 monozygotic and 372 same-sex dizygotic twin pairs (1,502 subjects total) from the Colorado Twin Registry in the United States (65%) and the National Health and Medical Research Council's Australian Twin Registry (35%). This unselected population sample is part of the International Longitudinal Twin Study (ILTS) of early reading and attention development (e.g., Samuelsson et al., 2005; Willcutt et al., 2007). This study is ongoing, and data is not currently available from Australia for the post-4th grade time point. At the time of the initial assessment with the SWAN, all subjects had just completed kindergarten, with ages ranging from 64 to 92 months (mean=75.7 months). Mean years of the subjects' parents' education was 14. Zygosity was confirmed with DNA collected via cheek swabs or, in a small number of cases, by selected items from the questionnaire by Nichols and Bilbro (1966).

Instruments

DBRS—The Disruptive Behavior Rating Scale is a caregiver-report questionnaire comprising 18 items targeting symptoms of inattention and HI, and 8 items measuring conduct disorder, the latter of which were not included in this study. The Inattention and HI items map onto the criteria for Attention Deficit/Hyperactivity Disorder as outlined by the Diagnostic and Statistical Manual, Fourth Edition (DSM-IV, American Psychiatric Association, 1994). The DBRS is scored on a 4-point Likert Scale, with anchors 0=*not at all*, 1=*sometimes*, 2=*often*, and 3=*very often*. The two-factor structure of the DBRS ADHD scales (i.e. inattention and hyperactivity/impulsivity) was established by DuPaul, Anastopoulos, et al. (1998) using a normative sample of 4,666 subjects ranging from 4 to 20 years old ($M=9.57$). Subsequent studies have further supported this structure in preschool aged children (e.g., Friedman-Weieneth et al., 2009). Adequate reliability, as well as convergent and discriminant validity of the DBRS have been established for both parent and teacher ratings (G. J. DuPaul et al., 1997; G. J. DuPaul, Power, McGoe, Ikeda, & Anastopoulos, 1998; Erford, 1997).

SWAN—The Strengths and Weaknesses of ADHD Symptoms and Normal Behavior scale is a more recently developed measure. The wording of the SWAN items differs from that of the DBRS in that the questions are phrased relative to normal behavior expectations, rather than with a particular focus on deficits. For example, item 3 on the SWAN reads, “Compared to other children, how well does this child...Listen when spoken to directly?” while the corresponding DBRS item reads, “Does not seem to listen when spoken to directly.” The SWAN is rated on a balanced 7-point Likert scale, with anchors of 0=*far below*, 1=*below*, 2=*slightly below*, 3=*average*, 4=*slightly above*, 5=*above*, and 6=*far above*. Despite frequent use of the SWAN in recent research (e.g., Cornish, Wilding, & Hollis, 2008; Parsons, Bowerly, Buckwalter, & Rizzo, 2007), there exist few independent studies of the measure’s psychometric properties, including incremental ability to predict above-average cognitive performance.

Interference Ratings—As part of the DBRS, parents also rated the degree to which symptoms of inattention and HI interfere with social interactions, academic performance, daily responsibilities, and recreational activities. These ratings were made on a 0 to 3 Likert scale identical to the DBRS, and were completed concurrently with DBRS ratings.

CTOPP—The Rapid Naming subtests from the Comprehensive Test of Phonological Processing (CTOPP: Wagner, Torgesen, & Rashotte, 1999) were used as measures of naming speed, which previous research has shown to load strongly on a general Processing Speed factor also defined by perceptual speed tasks that do not require naming (McGrath et al., 2010). Subtests administered at each time point were chosen according to developmental level. At post-Kindergarten, subjects completed two Rapid Color Naming, two Rapid Digit Naming, and two Rapid Letter Naming tests. At post-1st and 4th grades, subjects completed two Rapid Digit and two Letter Naming subtests. Raw end times for individual subtests were log transformed and summed to create Rapid Naming composites, where higher values reflected worse performance.

Procedure

Procedures for the international Longitudinal Twin Study have been described in detail elsewhere (Samuelsson et al., 2005; Willcutt et al., 2007), so the following description will focus only on the assessments relevant to this article. Subjects were assessed using rapid naming measures at the end of kindergarten and 1st grade in the Australian and US samples, and also at the end of 4th grade in the US sample. One-hour testing sessions were performed at home or at school in the US, and at school in Australia. Parent ratings of ADHD symptoms were obtained at K, 1st, 2nd and 4th (US only) grades while individual testing was completed in the US, and via mail in Australia.

Analyses were performed with SPSS 18.0 and MPlus 6.0. Inattention, Hyperactive-Impulsivity (HI), and Total ADHD composite scores were created, consistent with procedures used in previous studies of similar rating scales (e.g., Pelham, Gnagy, Greenslade, & Milich, 1992; Willcutt et al., 2007). On both measures, items 1–9 are intended to correspond with symptoms of Inattention; items 10–18 tap symptoms of HI. SWAN items were reverse-scored (i.e. *far below*=7 and *far above*=1) to match the directionality of DBRS symptomatic scores.

Results

Internal Validity

Internal consistency was estimated for both measures by calculating Cronbach’s alpha for scale and total scores across the four time points: post-K, post-1st, post-2nd and post-4th

grades. The DBRS scales were found to demonstrate positively skewed distributions at all test occasions, with the majority of participants receiving low scores, consistent with previous research. Hence, these values were logarithm transformed prior to analyses (Willcutt et al., 2007). Mean Cronbach's alpha for the SWAN scales was nearly identical to the DBRS ($\alpha=.88$).

Low-bound estimates of test-retest reliability for SWAN and DBRS subscales were calculated by correlating scale scores at each time point for MZ twins. As described in previous literature (Samuelsson et al., 2005), these coefficients are underestimates of the true reliability due to nonshared environmental factors that create differences between individuals in the twin pair. Nonetheless, both measures demonstrated good reliability, although the mean SWAN coefficient was significantly higher than the mean DBRS coefficient ($p<.01$). Reliability for the SWAN ranged from .72–.90 ($M=.82$); DBRS coefficients ranged from .58–.81 ($M=.71$). The Inattention scale showed significantly lower test-retest reliability than the other scales for both the SWAN and DBRS at all time points except post-4th grade. Reliability did not vary significantly by timepoint for the SWAN; however, the post-4th grade reliability coefficients were significantly lower than those at post-K for the DBRS.

Longitudinal stability was calculated for both measures by correlating scale and total scores across the four time points: post-K, post-1st, post-2nd and post-4th grades. Pearson product moment correlation coefficients for the SWAN scales were acceptable and comparable to DBRS coefficients at all time intervals, and ranged from $r=.57$ to $r=.75$ ($M=.67$), $p<.001$; DBRS coefficients ranged from $r=.51$ to $r=.76$ ($M=.66$), $p<.001$. The DBRS coefficients were slightly lower than values previously published by DuPaul, et al. (1998), who found Inattention $r=.78$, and HI $r=.86$, and Total $r=.85$. However, the time interval between reports in the DuPaul, et al. study was only four weeks, while the current study averaged longitudinal stabilities across grades K, 1, 2, and 4. To our knowledge, there are no published longitudinal stability values for the SWAN.

Factor analysis was performed using SPSS 18.0. A two factor, maximum-likelihood extraction model with Varimax rotation was tested on SWAN and DBRS scores, with a random selection of one twin from each pair at post-2nd grade ($N=532$). Post-2nd grade data was used because we expect symptoms of ADHD to become more stable with age, and because only US data was available at post-4th grade. Overall, the items loaded as expected for both measures, with items 1–9 loading together as an Inattention factor, and items 10–18 loading onto the Hyperactivity-Impulsivity factor (see Table 1). Items 10 and 11 (*fidgets* and *leaves seat in classroom*) cross-loaded in both models. The SWAN showed additional cross-loading of items 3 (*listens when spoken to*) and 16 (*reflects on questions, controls blurting out answer*). On average, SWAN factor loadings were higher than DBRS loadings (SWAN $M=.74$; DBRS $M=.63$). The higher loading of SWAN items on the two factors would be consistent with it having a more differentiated Likert scale (7 points, versus 4 on the DBRS). Similarly, the two factor model accounted for 74% of the variance in the SWAN, and 64% of the variance in the DBRS. However, the DBRS showed a better fit with this model ($\chi^2 = 490$, $df=118$) compared to the SWAN ($\chi^2 = 866$, $df=118$).

To test the difference in fit of the 2-factor structure, we used MPlus Version 6.0 to run a confirmatory factor analysis with the default geomin rotation. Maximum likelihood factor analysis was performed, and the Inattention and HI scales were allowed to correlate within measures. As in the exploratory Varimax rotation, the factor loadings on both the Inattention and HI factors were significantly higher on the SWAN as compared to the DBRS (see Table 1). Constraining the path estimates of corresponding items to be equal across measures resulted in a significantly worse fit for both scales (Inattention: $\Delta\chi^2=45.83$, $\Delta df=8$, $p<.001$;

HI: $\Delta\chi^2=64.57$, $\Delta df=8$, $p<.001$), confirming that the measures do not show weak factorial invariance. These results suggest that the variance at the favorable end of the SWAN item scores is being used in a meaningful way.

Histogram plots revealed evident skew and kurtosis for the DBRS scales (Inattention skew: .95–1.13, kurtosis: .74–1.70; HI skew: 1.37–1.56, kurtosis: 2.25–2.55). SWAN plots were within the normal range (Inattention skew: $-.29$ – $.09$, kurtosis: $.05$ – $.65$; HI skew: $-.33$ – $.04$, kurtosis: $-.25$ – $.55$). A bivariate scatterplot of SWAN and DBRS Total scale scores at Time 1 is shown in Figure 1, and is representative of plots at all four time points. As we would expect, the shape of the scatterplot demonstrates heteroscedasticity, meaning that there is more variance around the regression line at the low end of the DBRS scale. Whereas the DBRS attributes scores of zero to subjects who are free of negative symptoms (x-axis), SWAN ratings of these same subjects capture variability among their scores (y-axis). A child who demonstrates good attention and regulation of activity level will have a summed Total scale score of zero on the DBRS. In contrast, this same child will receive scores between 1–4 (in our reverse-scored dataset) for each item on the SWAN, and will thus have a Total scale score between 18 and 72. The graph shows us that the SWAN is more sensitive to individual differences at the good ends of the symptom phenotypes than is the DBRS.

To further test this point, we compared MZ and DZ correlations on ADHD subscale scores for probands scoring at the favorable versus unfavorable halves of the distributions (determined by a median split). Because we observe greater variance in the SWAN than the DBRS at the favorable ends of the ADHD dimensions, we likewise expected to find greater differences between the MZ and DZ correlations in the SWAN in the favorable halves. In contrast, we expected comparable differences between the MZ and DZ correlations for the unfavorable halves. However, the results showed that at both the favorable and symptomatic halves of the dimensions, the SWAN had greater and more significant differences (Table 2).

Convergent Validity

Currently, measures like the DBRS, which map directly onto the DSM-IV criteria for ADHD, are considered the gold standard in clinical assessment of ADHD. Thus, it is useful to test whether the SWAN scale scores are related to the DBRS scores. The DBRS scale distributions violated assumptions of normality, and the logarithm transformed DBRS scale scores exhibited a curvilinear relationship with the SWAN scale scores. Thus, rather than performing linear correlations, we first performed regression analyses using the log transformed DBRS values and a quadratic variable, $(\log_{10}\text{DBRS})^2$, to determine the amount of variance shared by the models, as measured by R (Table 3). The SWAN and DBRS share a moderate amount of variance ($MR=.57$).

Both the SWAN and DBRS scales are ordered categorical variables, and as such do not have a defined metric. Therefore, in order to estimate the degree of correlation between the two measures, we used a measure of association (Cramer's V) that is analogous to a Pearson correlation but does not assume a linear relationship. Like a correlation, Cramer's V ranges from -1 to 1 , with a value of zero indicating no association, and a value of 1 indicating a perfect association. In order to calculate Cramer's V, we first computed the chi-square for the difference between the expected number of individuals at each level of the SWAN and DBRS and the observed number of individuals in a contingency table. Then we entered that value into Equation 1.

$$\text{Cramer's } V = \sqrt{\chi^2 / (n(k-1))} \quad \text{Equation 1}$$

where n equals the number of participants and k equals the number of rows or columns, whichever is smaller.

Cramer's V statistics were calculated at concurrent time points to estimate the association between the two measures (Table 3). Total scale scores consistently showed the strongest relationship (M Cramer's $V = .53$), and both Inattention and HI scales demonstrated moderate cross-measure associations ($M = .48$).

Next, we truncated the SWAN item-level scores to match the DBRS scales. So, using our reverse-scored dataset, a SWAN score of 1–4 was coded as a 0 in the truncated variable, 5 was coded as 1, 6 as 2, and 7 as 3. Because both scales now violated the assumption of normality, we again used Cramer's V to estimate the agreement between the composite scale scores at each time point. Cramer's V values were similar to those achieved using the non-truncated SWAN scores: M Total $V = .56$, M Inattention $V = .50$, M HI $V = .50$.

Between-measure diagnostic agreement for ADHD subtypes was also calculated using Cramer's V statistic. In accordance with DSM-IV diagnostic guidelines, subjects were classified as ADHD-Inattentive Type if 6 or more items on the Inattentive scale were endorsed; Hyperactive Type if 6 or more HI items were endorsed; and Combined Type if 6 or more items on each scale were endorsed. An item was considered "endorsed" on the DBRS if it was scored as a 2 (*often*) or 3 (*very often*). As mentioned previously, SWAN scores were reversed to match directionality of the DBRS. On the SWAN, items were considered endorsed for scores of 5 (*slightly below*), 6 (*below*) or 7 (*far below*). We found slightly improved categorical agreement when we counted these three ratings as endorsements on the SWAN, as opposed to only counting the two highest ratings, as is done on the DBRS. Cramer's V statistics for categorical comparisons across measures were similar to the continuous values: M Inattentive = .34, M Hyperactive-Impulsive = .41, M Combined = .45 (Table 3). Cramer's V agreement for any ADHD diagnosis was .50.

Another way to examine categorical correlations is to calculate sensitivity and specificity for the SWAN, using the DBRS as the gold standard. For detection of all ADHD subtypes, the SWAN showed moderate sensitivity ($M = 58\%$) and excellent specificity ($M = 98\%$). However, it should be noted that given the imperfect reliability of the DBRS in this sample (categorical test-retest reliability $\sim .60$), it is both a mathematical and theoretical fallacy to assume that the DBRS is in fact the "gold standard." Error in both measures will lower the apparent sensitivity of the SWAN, and with this in mind we consider the observed sensitivity values to be acceptable, overall. Table 4 reports the number of individual cases identified by both measures, the SWAN only, and DBRS only at each time point. Overall, the SWAN identified more cases than did the DBRS, and raw agreement rates were comparable at each time point.

Heritability

It is extensively documented that symptoms of Inattention and HI are heritable, with one meta-analysis of 20 twin studies finding a mean heritability of .76 for total ADHD scores (Faraone, Spencer, Aleardi, Pagano, & Biederman, 2004). Thus, testing whether SWAN scores are heritable and genetically correlated with DBRS scores are two important convergent validity tests. We utilized univariate ACE models to estimate genetic (A), shared environment (C), and non shared environment (E) factors for DBRS and SWAN separately for each grade level using Mx (Neale, Boker, Xie, & Maes, 2006). Likelihood based 95% confidence intervals were used to estimate the significance of genetic and environmental estimates, with the inclusion of zero signifying a nonsignificant factor. Nonsignificant paths were then dropped from the model, and difference in chi-squares was used to confirm a nonsignificant loss in model fit. As before, the DBRS measures were log transformed to

minimize skew. The effects of age (age and age squared) and gender were regressed out prior to genetic analyses and the residuals were used.

We utilized an ACE Correlated Factors Model, although statistically equivalent to a Cholesky decomposition, the latent A C and E factors in this model are linked by a correlation r rather than common paths as in a Cholesky decomposition. This model was used to investigate the degree to which shared risk factors account for the comorbidity between SWAN and DBRS. The correlation r measures the degree of shared risk. Correlations can range from -1 to 1 , indicating the extent to which the same genes influence both traits irrespective of their specific heritabilities. Similarly, this model allowed us to test the extent to which shared and nonshared environmental factors are correlated between these measures. Nonsignificant factors were dropped from the model, with the difference in chi-squares was measured to confirm a nonsignificant loss in model fit. The results from a reduced AE model are presented in Table 5 because shared environment could be dropped without a significant loss in fit cross all grades and measures. Heritabilities for SWAN were similar across subtypes and grade levels, ranging from .69 to .89. Moreover, these estimates are similar to previous findings for ADHD scores.

Univariate A and E estimates for the DBRS from best fitting models at each grade level are presented in Table 5. Heritabilities were similar across subtypes and grade levels, ranging from .52 to .79. These estimates are similar to findings for SWAN scores.

Bivariate heritabilities (h^2_{xy}) are presented at the bottom of Table 5, which were moderate across time and type (i.e. Inattention, Hyperactivity-Impulsivity), ranging from .48 to .60. Nonshared environment correlations were moderate to large across time and subscale, ranging from .42 to .72.

Finally, we calculated the proportion of variance in the phenotypic correlations due to shared genetic factors (pg) across grade levels. The formula for the pg correlations for the SWAN and DBRS scales is:

$$pg_{xy} = h^2_{xy} / p_{xy} \quad \text{Equation 2}$$

where h^2_{xy} represents the bivariate heritability for the SWAN (x) and DBRS (y), and p_{xy} represents the phenotypic correlation between the two measures. Pg averaged across time points ranged from .78 to .90 for each subscale. The results show that genetic factors shared by the SWAN and DBRS account for a large amount of variance in the phenotypic correlation of the two measures.

External Validity

Of primary interest in this study is whether the SWAN offers improved ability over DSM-IV based ADHD measures to measure the full range of behavior, including performance at the good end of the two ADHD symptom dimensions. Thus, our analyses of external validity focused on testing whether the SWAN offers incremental external validity over the DBRS for predicting good cognitive performance and less behavioral interference.

To answer these questions, we performed linear regressions to determine whether the SWAN could predict cognitive performance and behavioral interference over and above the DBRS. We tested these regressions both predictively and concurrently. Using the “step” option in SPSS 6.0, DBRS Inattention was entered first, followed by DBRS HI, and then each SWAN scale, one at a time. Table 6 shows the results of Rapid Naming scores at Post-1st grade regressed on the post-K ADHD scales. Adding the SWAN Inattention scale resulted in a significant R^2 change (.068), and Rapid Naming scores were best predicted by

SWAN Inattention scale scores ($\beta = .34$) when all predictors were included in the model. The results of Rapid Naming regressions at post-4th, (where only data from the US is available) were similar and therefore not reported in detail here.

In contrast, hierarchical regressions of Post-2nd grade Interference scores on ADHD subscales revealed that the SWAN scales explained very little variance in behavioral functioning when DBRS scores had already been entered ($\Delta R^2 = .005$). We chose to use post-2nd grade data for this analysis as it represented the longest, more stringent predictive time interval where both US and Australia data was available. The DBRS HI scale was the best predictor of behavioral functioning ($\beta = .209$) when all four variables were included in the model (Table 6).

Finally, post-4th grade RAN scores were regressed onto DBRS and truncated SWAN scale scores separately, to compare the amount of variance accounted for by each measure. The truncated SWAN scale scores accounted for more variance than did the DBRS scores, with a mean R^2 value of .056 for the SWAN total score, and .037 for the DBRS.

Discussion

We examined the psychometric properties of a new measure of ADHD symptoms, the Strengths and Weaknesses of ADHD Symptoms and Normal Behavior (SWAN) scale. The SWAN demonstrates internal validity comparable to the DBRS, as measured by internal consistency, longitudinal stability and factor structure. The similar factor structures also provide evidence of convergent validity between the SWAN and DBRS, in that the two questionnaires are measuring overlapping constructs. Further, the significantly higher factor loadings on the SWAN indicate that the variance measured by this scale at the good ends of Inattention and HI is related to the original two factors, and thus meaningful.

All tests of convergent validity indicate that the SWAN and DBRS are measuring similar constructs. Linear regressions established a moderate amount of shared variance between the SWAN and DBRS subscales. Continuous and categorical agreement between the measures were estimated using Cramer's V statistic, and found to be moderate. One possible confound in comparing the measures is that the SWAN has a longer Likert scale. With extended Likert scales, reporters may be less likely to endorse extreme scores. However, Table 4 demonstrates that on average, the SWAN identified more cases than did the DBRS, so it does not appear that this potential confound was a factor in this sample. Although agreement between the two measures is not perfect, we expected the SWAN to identify different cases than the DBRS because the measures are in fact partly independent, and both are further limited by their own imperfect reliabilities. Further, categorical correlations are always expected to be lower than continuous associations. Taken together, these tests support the use of the SWAN for clinical purposes as well as research in the diagnosis of ADHD subtypes.

Bivariate heritability analyses demonstrated that shared genetic influences account for a large proportion of the phenotypic correlation between the measures. On the other hand, the genetic correlation is lower than we expected, and it is not due to a difference in genetic influences specifically at the favorable ends of the scales. Rather, the results suggest that the SWAN is measuring some genetic factors that are distinct from those measured by the DBRS on both ends of the dimensions. This supports the SWAN as a distinct and useful measure of variance not captured by the DBRS.

Our primary question in this study was whether the SWAN has additional information about children performing at the favorable ends of Inattention and HI. As predicted, the SWAN does have variance at the positive ends of Inattention, HI and Combined subscales, as

demonstrated by heteroscedasticity with the DBRS scale. Further, hierarchical regressions of a measure of cognitive functioning showed the SWAN Inattention scale to be the only independent predictor of cognitive performance, indicating that this subscale in particular is capturing meaningful variance over and above those of the DBRS. Finally, the differences between MZ and DZ correlations were greater for the SWAN at both ends of the symptom dimensions, suggesting that the SWAN offers increased variance for the full spectrum of behaviors.

However, the SWAN did not offer incremental validity for prediction of interference. This may be related to two limitations: 1) the Interference ratings were collected at the same time as the DBRS ratings, thus they may be biased to be more closely related to the DBRS scores; and 2) the Interference ratings were made on a 0–3 Likert scale, with no variance on the “good” end of interference, so the SWAN’s ability to capture variance at this good end was likely not maximized during these analyses.

Research to date has struggled to identify consistently strong cognitive correlates of ADHD (Nigg, Willcutt, Doyle, & Sonuga-Barke, 2005). The current study suggests that skewed measures of ADHD symptomology, such as the DBRS, may be inadvertently limiting the potential for those correlations. In contrast, the SWAN provides sufficient variance at both ends of the Inattention and HI spectrums to capture a statistical relationship. Thus, future investigations of candidate cognitive endophenotypes of ADHD would benefit from use of the SWAN as the behavioral measure.

Overall, the SWAN appears to be a useful tool for examining the etiology of the two ADHD symptom dimensions, both at the adaptive and impaired ends of performance. Given the increased range and complexity of SWAN scores, future work could include development of normative scores for diagnosis of ADHD. This would be an improvement over the symptom-count method employed by the DBRS and other DSM-IV-based measures for prediction of cognitive as well as behavioral functioning. Normative scores would also serve to strengthen the SWAN as a research tool, and would provide a rich set of data with which to examine candidate cognitive endophenotypes for adaptive functioning correlates of development of Inattention and HI.

References

- Achenbach, TM. Manual for the child behavior Checklist/4–18 and 1991 profile. University of Vermont, Burlington, VT: University of Vermont Department of Psychiatry; 1991.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, D.C: American Psychiatric Association; 1994.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, D.C: American Psychiatric Association; 2000. -TR
- Barkley, R.; Murphy, K. Attention deficit hyperactivity disorder: A clinical workbook. 2. New York: Guilford Press; 1998.
- Bussing R, Fernandez M, Harwood M, Hou W, Garvan CW, Eyberg SM, Swanson JM. Parent and teacher SNAP-IV ratings of attention deficit hyperactivity disorder symptoms: Psychometric properties and normative ratings from a school district sample. *Assessment* (Odessa, Fla). 2008; 15(3):317.
- Center for Disease Control. Mental health in the united states: Prevalence of diagnosis and medication treatment for attention-Deficit/Hyperactivity disorder -united states, 2003. *Morbidity and Mortality Weekly Report*. 2005; 54(34):842–847. [PubMed: 16138075]
- Cornish KM, Wilding JM, Hollis C. Visual search performance in children rated as good or poor attenders: The differential impact of DAT1 genotype, IQ, and chronological age. *Neuropsychology*. 2008; 22:217–225. [PubMed: 18331164]

- DuPaul GJ, Anastopoulos AD, Power TJ, Reid R, Ikeda MJ, McGoey KE. Parent ratings of attention deficit/hyperactivity disorder symptoms: Factor structure and normative data. *Journal of Psychopathology and Behavioral Assessment*. 1998; 20:83–102.
- DuPaul GJ, Power TJ, Anastopoulos AD, Reid R, McGoey KE, Ikeda MJ. Teacher ratings of attention deficit/hyperactivity disorder symptoms: Factor structure, normative data, and psychometric properties. *Psychological Assessment*. 1997; 9:436–444.
- DuPaul GJ, Power TJ, McGoey KE, Ikeda MJ, Anastopoulos AD. Reliability and validity of parent and teacher ratings of attention-deficit/hyperactivity disorder symptoms. *Journal of Psychoeducational Assessment*. 1998; 16:55–68.
- DuPaul GJ, Power TJ, McGoey KE, Ikeda MJ, Anastopoulos AD. Parent ratings of attention-deficit/hyperactivity disorder symptoms: Factor structure and normative data. *Journal of Psychopathology and Behavioral Assessment*. 1998; 20(1):83–102.
- Erford BT. Reliability and validity of scores on the disruptive behavior rating scale - teacher version (DBRS-T). *Educational and Psychological Measurement*. 1997; 57(2):329.
- Faraone SV, Spencer T, Aleardi M, Pagano C, Biederman J. Meta-analysis of the efficacy of methylphenidate for treating adult attention-deficit/hyperactivity disorder. *Journal of Child Psychopharmacology*. 2004; 24(1):24–29.
- Friedman-Weieneth J, Doctoroff GL, Harvey EA, Goldstein LH. The disruptive behavior rating Scale —Parent version (DBRS-PV): Factor analytic structure and validity among young preschool children. *Journal of Attention Disorders*. 2009; 13(1):42. [PubMed: 18753403]
- Hay D, Bennett KS, Levy F, Sergeant J, Swanson J. A twin study of attention-Deficit/Hyperactivity disorder dimensions rated by the strengths and weaknesses of ADHD-symptoms and normal-behavior. *Biological Psychiatry* (1969). 2007; 61(5):700.
- McGrath, L.; Shanahan, M.; Santerre-Lemmon, L.; Barnard, H.; Willcutt, E.; Olson, R.; Pennington, B. Unpublished Manuscript. 2010. A multiple deficit model of reading disability and attention-Deficit/Hyperactivity disorder: Searching for a shared cognitive deficit.
- Neale, MC.; Boker, SM.; Xie, G.; Maes, HH. Mx: Statistical modeling manual. Richmond, VA: Department of Psychiatry, Medical College of Virginia; 2006.
- Nichols RC, Bilbro WC. The diagnosis of twin zygosity. *Acta Genetica*. 1966; 16:265–275.
- Nigg J, Willcutt EG, Doyle AE, Sonuga-Barke EJ. Causal heterogeneity in attention-deficit/hyperactivity disorder: Do we need neuropsychologically impaired subtypes? *Biological Psychiatry*. 2005; 57(11):1224. [PubMed: 15949992]
- Parsons TD, Bowerly T, Buckwalter JG, Rizzo A. A controlled clinical comparison of attention performance in children with ADHD in a virtual reality classroom compared to standard neuropsychological methods. *Child Neuropsychology*. 2007; 3:363–381. [PubMed: 17564852]
- Pelham WE, Gnagy EM, Greenslade KE, Milich R. Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1992; 31:210–218. [PubMed: 1564021]
- Pennington, BF. The development of psychopathology: Nature and nurture. New York, NY: Guilford Press; 2002.
- Polderman TJC, Derks EM, Hudziak JJ, Verhulst FC, Posthuma D, Boomsma DI. Across the continuum of attention skills: A twin study of the SWAN ADHD rating scale. *Journal of Child Psychology and Psychiatry*. 2007; 48(11):1080. [PubMed: 17995483]
- Samuelsson S, Byrne B, Quain P, Wadsworth S, Corley R, DeFries JC, Olson R. Environmental and genetic influences on prereading skills in australia, scandinavia, and the united states. *Journal of Educational Psychology*. 2005; 97(4):705–722.
- Stevenson J, Asherson P, Hay D, Levy F, Swanson J, Thapar A, Willcutt E. Characterizing the ADHD phenotype for genetic studies. *Developmental Science*. 2005; 8(2):115–121. [PubMed: 15720369]
- Swanson, J.; Shuck, S.; Mann, M.; Carlson, C.; Hartman, K.; Sergeant, J.; McCleary, R. Categorical and dimensional definitions and evaluations of symptoms of ADHD: The SNAP and SWAN rating scales. Irvine: University of California, Irvine; 2006.
- Wagner, RK.; Torgesen, JK.; Rashotte, CA. The comprehensive test of phonological processes (CTOPP). Austin, TX: PRO-ED; 1999.

- Willcutt EG, Betjemann RS, Wadsworth SJ, Samuelsson S, Corley R, DeFries JC, Olson RK. Preschool twin study of the relation between attention-deficit/hyperactivity disorder and prereading skills. *Reading and Writing*. 2007; 20:103–125.
- Young D, Levy F, Martin NC, Hay DA. Attention deficit hyperactivity disorder: A rasch analysis of the SWAN rating scale. *Child Psychiatry and Human Development*. 2009; 40(4):543. [PubMed: 19455417]

\$watermark-text

\$watermark-text

\$watermark-text

Figure 1.
Heteroscedasticity of SWAN and DBRS Total Scores at Post-K

Table 1

Factor Loadings for Two-Factor Varimax Rotation Model

	DBRS		SWAN	
	Inattention	HI	Inattention	HI
1. Attention to details	.639	.103	.743	.296
2. Sustaining attention	.675	.265	.766	.391
3. Listening when spoken to	.543	.381	.664	.495
4. Following through on instructions	.750	.171	.735	.371
5. Organizing tasks	.732	.183	.805	.287
6. Avoiding tasks requiring mental effort	.623	.094	.721	.290
7. Losing things necessary for tasks	.555	.257	.800	.291
8. Distractibility	.707	.294	.675	.375
9. Forgetfulness	.676	.256	.766	.277
10. Fidgeting	.436	.502	.433	.711
11. Remaining seated in class	.426	.468	.443	.721
12. Running, climbing	.310	.517	.345	.684
13. Playing quietly	.303	.621	.285	.829
14. On the go	.230	.619	.316	.853
15. Excessive talking	.057	.656	.192	.806
16. Waiting to answer questions	.089	.654	.437	.552
17. Waiting turn	.209	.717	.377	.734
18. Interrupting, intruding on others	.216	.701	.347	.688

Table 2

Differences Between MZ and DZ Correlations by Median Split for ADHD Scales

		SWAN $r_{MZ-r_{DZ}}$	DBRS $r_{MZ-r_{DZ}}$
Post-K	Good Attention	0.492	0.432
	Bad Attention	0.290	0.221
	Good HI	0.283	0.293
	Bad HI	0.421	0.338
	Good Total	0.412	0.340
	Bad Total	0.309	0.214
Post-1 st	Good Attention	0.420	0.327
	Bad Attention	0.479	0.362
	Good HI	0.279	0.098
	Bad HI	0.457	0.503
	Good Total	0.403	0.231
	Bad Total	0.464	0.395
Post-2 nd	Good Attention	0.312	0.383
	Bad Attention	0.470	0.295
	Good HI	0.232	0.128
	Bad HI	0.772	0.050
	Good Total	0.245	0.469
	Bad Total	0.712	0.319
Post-4 th	Good Attention	0.438	0.351
	Bad Attention	0.516	0.291
	Good HI	0.336	0.249
	Bad HI	0.437	0.354
	Good Total	0.454	0.094
	Bad Total	0.554	0.093
<i>Mean</i>	Good Attention	0.416	0.373
	Bad Attention	0.439	0.292
	Good HI	0.283	0.192
	Bad HI	0.522	0.311
	Good Total	0.379	0.284
	Bad Total	0.510	0.255

Table 3

SWAN and DBRS Scale and Diagnostic Agreement

		Inattention	HI	Combined
Post-K	R	0.607	0.481	0.593
	Cramer's V Scale Agreement	0.516	0.475	0.502
	Cramer's V Diagnostic Agreement	0.379	0.435	0.563
Post-1 st	R	0.624	0.476	0.607
	Cramer's V Scale Agreement	0.461	0.475	0.512
	Cramer's V Diagnostic Agreement	0.433	0.259	0.427
Post-2 nd	R	0.657	0.475	0.617
	Cramer's V Scale Agreement	0.471	0.509	0.521
	Cramer's V Diagnostic Agreement	0.297	0.37	0.367
Post-4 th	R	0.649	0.454	0.597
	Cramer's V Scale Agreement	0.486	0.469	0.578
	Cramer's V Diagnostic Agreement	0.256	0.563	0.456

\$watermark-text

\$watermark-text

\$watermark-text

Table 4
Case Agreement Between SWAN and DBRS as a Function of Subtype and Grade Level

		ADHD			
		ADHD-I	ADHD-HI	ADHD-C	Any ADHD
Post-K	Both	4	7	3	16
	SWAN Only	11	8	4	21
	DBRS Only	3	9	1	11
Post-1 st	Both	6	3	4	18
	SWAN Only	12	12	10	29
	DBRS Only	4	5	2	6
Post-2 nd	Both	5	5	3	22
	SWAN Only	14	7	13	25
	DBRS Only	8	9	1	9
Post-4 th	Both	4	4	3	16
	SWAN Only	18	3	11	27
	DBRS Only	5	3	0	3
Mean	Both	5	5	3	18
	SWAN Only	14	8	10	26
	DBRS Only	5	7	1	7

Table 5

Univariate estimates of A and E and Genetic and Environmental Correlations as a Function of Grade Level

SWAN Univariate values				
Grade	Component	Inattention	Hyperactivity-Impulsivity	Total Score
Post-K	A	.75 (.69 – .79)	.84 (.80 – .87)	.83 (.79 – .86)
	E	.25 (.21 – .32)	.16 (.13 – .20)	.17 (.14 – .21)
Post-1 st	A	.80 (.75 – .84)	.89 (.87 – .91)	.89 (.87 – .91)
	E	.20 (.16 – .25)	.11 (.09 – .13)	.11 (.09 – .13)
Post-2 nd	A	.77 (.72 – .81)	.89 (.86 – .91)	.85 (.82 – .88)
	E	.23 (.19 – .28)	.11 (.09 – .14)	.15 (.12 – .19)
Post-4 th	A	.69 (.59 – .76)	.86 (.82 – .89)	.82 (.77 – .86)
	E	.31 (.24 – .41)	.14 (.11 – .18)	.18 (.14 – .28)
DBRS Univariate Values				
Grade	Component	Inattention	Hyperactivity-Impulsivity	Total Score
Post-K	A	.71 (.65 – .76)	.79 (.74 – .82)	.81 (.78 – .84)
	E	.29 (.24 – .35)	.21 (.18 – .26)	.19 (.16 – .22)
Post-1 st	A	.64 (.60 – .70)	.79 (.75 – .83)	.77 (.72 – .81)
	E	.36 (.30 – .43)	.21 (.17 – .25)	.23 (.19 – .28)
Post-2 nd	A	.57 (.49 – .64)	.77 (.71 – .81)	.71 (.65 – .76)
	E	.43 (.36 – .52)	.23 (.19 – .29)	.29 (.24 – .35)
Post-4 th	A	.52 (.39 – .61)	.70 (.63 – .76)	.65 (.57 – .72)
	E	.49 (.39 – .61)	.30 (.24 – .37)	.35 (.28 – .43)
SWAN and DBRS Common Variance				
Grade	Component	Inattention	Hyperactivity-Impulsivity	Total Score
Post-K	A	.53 (.47 – .59)	.56 (.49 – .62)	.54 (.48 – .56)
	E	.50 (.41 – .58)	.66 (.58 – .72)	.61 (.53 – .68)
Post-1 st	A	.56 (.50 – .61)	.60 (.53 – .65)	.56 (.57 – .71)
	E	.42 (.32 – .52)	.67 (.60 – .74)	.59 (.50 – .66)
Post-2 nd	A	.50 (.43 – .56)	.60 (.53 – .66)	.53 (.46 – .59)
	E	.57 (.47 – .65)	.69 (.62 – .75)	.65 (.57 – .71)
Post-4 th	A	.48 (.40 – .55)	.54 (.43 – .62)	.49 (.40 – .57)
	E	.64 (.53 – .72)	.72 (.67 – .80)	.73 (.65 – .79)

Note. 95% confidence interval in parentheses. Across all measures and time points, shared environment in the univariate models could be dropped from the model without a significant loss in fit.

Table 6

Hierarchical Regressions of External Validators on Post-K ADHD Subscales

Model: Post-1st Rapid Naming on Post-K ADHD					
		Standardized Beta	t	p	R
1	DBRS Inattention	.234	5.597	.000	.234
2	DBRS Inattention	.274	5.513	.000	
	DBRS HI	-.074	-1.481	.139	.242
3	DBRS Inattention	.075	1.312	.190	
	DBRS HI	-.060	-1.255	.210	
	SWAN Inattention	.324	6.472	.000	.356
4	DBRS Inattention	.070	1.179	.239	
	DBRS HI	-.049	-.825	.410	
	SWAN Inattention	.340	4.961	.000	
	SWAN HI	-.023	-.340	.734	.356
Model: Post-2nd Grade Interference on Post-K ADHD					
		Standardized Beta	t	p	R
1	DBRS Inattention	.265	4.876	.000	.265
2	DBRS Inattention	.117	1.848	.066	
	DBRS HI	.266	4.198	.000	.345
3	DBRS Inattention	.083	1.077	.282	
	DBRS HI	.270	4.238	.000	
	SWAN Inattention	.051	.760	.448	.348
4	DBRS Inattention	.115	1.419	.157	
	DBRS HI	.209	2.687	.008	
	SWAN Inattention	-.031	-.341	.733	
	SWAN HI	.119	1.338	.182	.355