ADHD Traits in German School-Aged Children: Validation of the German Strengths and Weaknesses of ADHS Symptoms and Normal Behavior (SWAN-DE) Scale

Yoanna Schulz-Zhecheva[1], Manuel Voelkle[2,3], André Beauducel[4], Nadja Buch[1], Christian Fleischhaker[1], Stephan Bender[5], Christopher W. N. Saville[6], Monica Biscaldi[1], and Christoph Klein[1,5,6]

Abstract

Objective: The Strengths and Weaknesses of ADHS–Symptoms and Normal Behavior (SWAN) Scale has been developed to support the dimensional assessment of ADHD symptoms by capturing variance on both poles of the ADHD continuum. The present study provides the first validation of the German version of SWAN (SWAN-DE). **Method:** Based on a sample of N_1 = 343 children from the general population and N_2 = 62 children with ADHD, both aged between 8 and 18 years, normality, internal consistency, test–retest reliability, and different validity indices were examined. **Results:** SWAN was characterized by normally distributed scores, good to excellent reliability, and factorial validity. It showed high diagnostic utility in discriminating between patients with ADHD and healthy controls and significant correlations to related clinical scales and neuropsychological constructs, such as intra-subject variability. **Conclusion:** The present study reveals the excellent psychometric properties of SWAN-DE, which can now be usefully applied in the German-speaking countries as well as in cross-national studies. *(J. of Att. Dis. 2019; Vol. 23(6) 553-562)*

Keywords

ADHD, behavioral assessment, rating scales, symptomatology, SWAN

Introduction

ADHD, one of the most common neurodevelopmental disorders, is characterized by symptoms of inattention, hyperactivity, and impulsivity. The conceptualization of ADHD as a dimensional rather than a categorical construct (Levy, Hay, McStephen, Wood, & Waldman, 1997) considers the ability to control attention and activity as continua reaching from excellent control to inattention and hyperactivity/impulsivity. Thus, the clinical features of persons with ADHD represent the negative pole of a dimension that is continuously distributed within the general population. Initial assessments of this dimensional model of ADHD in the general population have specified a bifactor model, consisting of a general ADHD factor complemented by two specific ADHD factors (inattention and hyperactivity/impulsivity; Normand, Flora, Toplak, & Tannock, 2012).

Although empirical evidence clearly supports the conceptualization of ADHD as a dimensional rather than a categorical disorder (Coghill & Sonuga-Barke, 2012), the categorical and deficit-oriented view still dominates the

clinical assessment of the disorder. Thus, in common clinical interviews, ADHD symptoms are categorically scored as present or absent, and even widely used clinical scales truncate the full range of behaviors by relying solely on deficit-oriented items and neglecting co-occurring strengths of patients (Swanson et al., 2006). This approach has at least three major disadvantages. First, deficit-oriented scales typically produce a skewed distribution in the general population, where most of the symptoms occur "never" or "sometimes," that is, the mean scores lie between 0 and

[1]Medical Faculty, University of Freiburg, Germany
[2]Institute of Psychology, Humboldt University of Berlin, Germany
[3]Max Planck Institute for Human Development, Berlin, Germany
[4]Institute of Psychology, University of Bonn, Germany
[5]Medical Faculty, University of Cologne, Germany
[6]School of Psychology, Bangor University, UK

Corresponding Author:
Christoph Klein, Department of Child and Adolescent Psychiatry, University of Freiburg, Hauptstrasse 8, D-79104 Freiburg, Germany.
Email: christoph.klein.kjp@uniklinik-freiburg.de

1. Such skewed distributions, however, can result in an overestimation of the prevalence of ADHD if cutoff scores based on mathematical transformations that assume a normal distribution of scores are used (Swanson et al., 2006). Second, due to the described heterogeneity of ADHD, it is to be expected that individuals who show deficits in some ADHD-related behaviors show strengths in others. Thus, by only assessing the negative pole of the continuum, individual strengths of clinical importance may be overlooked. Third, the skewed distribution as a result of not assessing meaningful variance on the positive end of the distribution can have dramatic methodological consequences in population-based studies on ADHD (e.g., genetic studies) by reducing their statistical power and validity.

To overcome these shortcomings of common ADHD scales, Swanson introduced the Strengths and Weaknesses of ADHD–Symptoms and Normal Behavior (SWAN) Scale (Swanson et al., 2006), a questionnaire with 18 items based on the *Diagnostic and Statistical Manual of Mental Disorders, Fourth/Fifth Edition* (*DSM-IV/V*) symptom criteria for ADHD (American Psychiatric Association [APA], 1994; APA, 2013). Due to the neutral wording of the items (e.g., "Remember daily activities" instead of "Often is forgetful in daily activities" etc.) and the well differentiating 7-point-scale, the SWAN Scale is the first diagnostic instrument assessing ADHD symptoms in a truly dimensional manner both at the positive and negative sides of the distribution.

Various validation studies have reported excellent psychometric properties of the original SWAN Scale, including normally distributed scores in the general population, high internal consistency and moderate retest reliability, consistent factorial structure, and high convergent validity as revealed by significant correlations between SWAN scores and widely applied diagnostic instruments (Arnett et al., 2013; Hay, Bennett, Levy, Sergeant, & Swanson, 2007; Lakes, Swanson, & Riggs, 2012; Polderman et al., 2007; Swanson et al., 2006). As a result, the SWAN Scale has been applied in a variety of studies, investigating the genetics (Hay et al., 2007; Polderman et al., 2007) and possible endophenotypes of ADHD (Crosbie et al., 2013). Not only has the SWAN Scale been confirmed as a useful instrument for assessing ADHD symptoms in the general population, but it also appears to show high sensitivity and specificity in differentiating patients with ADHD from healthy controls (Lai et al., 2011; Robaey, Amre, Schachar, & Simard, 2007) and other psychiatric disorders (Chan, Lai, Luk, Hung, & Leung, 2014).

While the original SWAN Scale has been translated to various languages such as Spanish (Lakes et al., 2012), French (Robaey et al., 2007), and Chinese (Lai et al., 2011), a translation to German is still missing. The purpose of the present study was therefore to introduce the German version of the SWAN Scale (SWAN-DE) by assessing its psychometric properties in school-aged German children with and without a diagnosis of ADHD, including internal consistency and test–retest reliability, factorial and concurrent validity with established clinical scales, and behavioral measures of processing speed and intra-subject variability (ISV; Kuntsi & Klein, 2011).

Method

This study was approved by the ethics committee of the University of Freiburg. Written informed consent was given by all participants or their parents.

Participants

Analyses were based on data from two partly overlapping samples of $N_1 = 343$ children from the general population (6% ADHD; 48% male; age: $M \pm SD = 12.79 \pm 2.87$; IQ: $M \pm SD = 108.9 \pm 14.65$) and $N_2 = 62$ children with ADHD (65% male; age: $M \pm SD = 12.77 \pm 2.22$; IQ: $M \pm SD = 93.96 \pm 13.32$), both aged between 8 and 18 years. The general population sample was created by adding 20 randomly selected children from the ADHD sample to a sample of 323 healthy participants without any present neurological or psychiatric disorders to ensure a representative prevalence of ADHD diagnosis in Germany of about 6% (Huss, Hölling, Kurth, & Schlack, 2008).

The parents of $N_3 = 77(22\%)$ participants from the general population, aged between 8 and 11 years (4% ADHD; 45% male; age: $M \pm SD = 10.42 \pm 1.05$; IQ: $M \pm SD = 110.8 \pm 15.04$), filled out the questionnaire for a second time approximately 27 weeks after their first participation. Four questionnaires were excluded from further analysis because they were completed less than 21 weeks or more than 33 weeks after the first participation.

Healthy participants were recruited through advertisements in schools. Of the ADHD sample, 55(89%) participants were outpatients of the Department of Child and Adolescent Psychiatry, who had received an ADHD diagnosis according to International Classification of Diseases-10 (ICD-10) criteria by experienced psychiatrists based on Parent and Teacher Rating Scales, medical files, and a clinical behavioral observation in a standardized group situation (e.g., test taking, structured play). The remaining seven patients had received an ADHD diagnosis from experienced child psychiatrists in private practices not affiliated with the Department of Child and Adolescent Psychiatry. Excluding these patients from the analysis did not affect the results. Methylphenidate medication was taken by 49 (79%) patients at the time of the study, but it was interrupted for at least 24 hr before participation. Of the 20 patients included in the general population sample, 19 (95%) were regularly taking methylphenidate medication.

Measures

SWAN (Swanson et al., 2006) is a dimensional rating scale, created through neutral rewording of the 18 symptom criteria of ADHD from *DSM-IV*. The rater is asked to compare the child with other children regarding 18 ADHD-related behaviors on a 7-point scale anchored to average behavior, reaching from *far below average* (0) to *far above average* (6). Thus, higher SWAN scores indicate less ADHD symptoms. Mean scores were computed for both subscales, Attention Deficit (SWAN-AD, nine items) and Hyperactivity/Impulsivity (SWAN-HI, nine items), as well as the total scale (SWAN-TOT).

The original SWAN was translated to German according to the guidelines for transcultural research proposed by Beaton, Bombardier, Guillemin, and Ferraz (2000), including various phases, such as (a) translation by two independent bilingual researchers, (b) back-translation by a third independent bilingual researcher, (c) group discussions, and (d) consultations with clinical experts.

The Child Behavior Checklist (CBCL; Achenbach, 1991) is a widely applied clinical parent report scale, assessing the frequency and intensity of cognitive, emotional, and behavioral problems in children on a 3-point scale, reaching from 0 (*not true*) to 2 (*very or often true*). The Attention Problem subscale (CBCL-AP) has 11 items and a maximum sum score of 22. Higher CBCL-AP scores indicate more ADHD symptoms.

"Fremdbeurteilungsbogen für ADHS" (FBB-ADHS; Döpfner, Görtz-Dorten, & Lehmkuhl, 2008) is a German ADHD rating scale based on ICD-10 and *DSM-IV* diagnostic criteria. It consists of 20 items regarding ADHD symptoms and additional items related to symptom onset, duration, functional impairment, and so on. The 4-point response scale ranges from 0 (*rarely*) to 3 (*very often*). Mean scores can be computed for the total scale (FBB-ADHS-TOT) as well as for the subscales: Inattention (FBB-ADHS-IN, nine items) and Hyperactivity/Impulsivity (FBB-ADHS-HI, nine items).

Processing speed (PS) and ISV were assessed as intra-individual mean reaction time and intra-individual reaction time standard deviation, respectively, from four parallel computer-based *choice reaction time tasks* (CRTs) adapted from Oberauer, Süß, Wilhelm, and Wittman (2003). In every CRT, 128 stimuli were presented successively in a pseudo-randomized order with an inter-stimulus interval of 2 s. The participants were asked to categorize the stimulus into one of two categories by pressing one of two response buttons on the keyboard as quickly and as accurately as possible. In CRT odd–even, participants were presented with two-digit numbers and were asked to categorize them as either odd or even. In CRT large–small, participants had to decide whether the number was larger or smaller than 50. In CRT up–down, the direction of arrows pointing upward or downward in different degrees was to be defined. In CRT above–below, sloping lines were presented either above or below the middle of the computer screen, and their position was to be indicated.

Procedure

SWAN was filled out by the parents of the healthy participants as a part of an online questionnaire, including CBCL and other clinical scales. Parents of participants with ADHD received a printed version and additionally filled out FBB-ADHS.

Participants were administered the four CRTs as part of a larger computer-based test battery, including working memory tasks, speed tasks, and an IQ testing. The results of these tasks will be presented elsewhere. The testing sessions took part in small groups of two to five participants, separated by partitions and wearing headphones with an active noise-cancellation system. At the beginning of every task, instructions were read to the participants, and 11 practice trials were administered. If more than two responses (18%) were errors, the practice trials were repeated. During the tests, the category–response button assignment was visible for the participants. Two patients with ADHD discontinued preterm their participation in the computer-based testing due to fatigue; therefore, they will be excluded from the analysis regarding PS and ISV.

Data Analysis

Basic data analyses were computed using the software package R (R Development Core Team, 2015). Mean scores were computed for the total scale and the subscales of SWAN, and the effects of gender and age were estimated for the population-based sample via univariate two-way ANOVAs using SPSS for Windows, Version 21. Because no major differences between the genders were evident, for the following analysis, data were pooled over male and female participants. Due to the large age range in the sample, linear age effects were partialled out from the data prior to further analysis. Results based on raw scores did not differ from results based on age-residualized scores though. For reasons of comparability to other studies with smaller age ranges, results based on age-residualized data will be reported, except for the identification of cutoff scores and the comparison between participants with and without ADHD. In addition, normality of the distribution and reliability indices (Cronbach's alpha, Pearson's correlation coefficients between the scores on the two measurement occasions) were examined. The factorial structure of SWAN was assessed using principal component analysis (PCA) with promax rotation and confirmatory factor analysis (CFA) with a bifactor
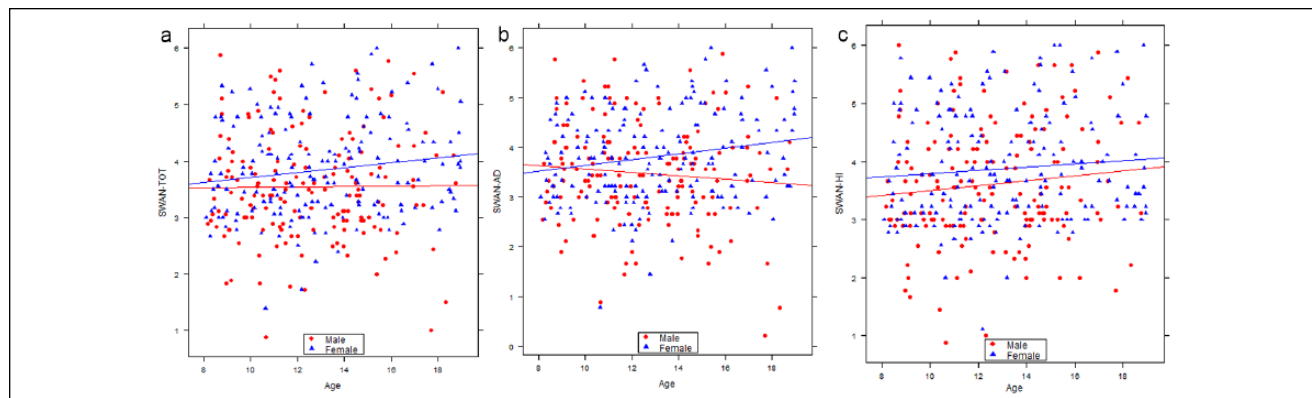
Figure 1. Relationships between SWAN-DE total score and subscale scores and age in male and female participants.
Note. SWAN-TOT = SWAN total score; SWAN-AD = SWAN subscale Attention Deficits; SWAN-HI = SWAN subscale Hyperactivity/Impulsivity; SWAN = Strengths and Weaknesses of ADHD–Symptoms and Normal Behavior; SWAN-DE = Strengths and Weaknesses of ADHS Symptoms and Normal Behavior, German Version

model assuming one general ADHD factor (G-ADHD) and two specific factors for inattention and hyperactivity/impulsivity (SP-AD and SP-HI). CFA was estimated in MPlus, Version 7.31 (Muthén & Muthén, 2010), using the Maximum likelihood estimation with robust standard errors (MLR) estimator. To estimate convergent validity, Pearson's correlations between SWAN and CBCL-AP as well as FBB-ADHS were computed. In addition, differences between patients and controls after controlling for age and gender effects were estimated using two-way ANCOVA. Receiver operating curve (ROC) analysis was computed via SPSS to determine the ability of SWAN to differentiate between participants with and without ADHD. Concurrent validity was further estimated based on the correlations between SWAN scores and ISV as well as PS in a structural equation model (SEM) using MPlus. Indicators for PS were intra-individual mean reaction times (M_i) from CRT odd–even and CRT up–down; indicators for ISV were intra-individual standard deviations of reaction times ($RTSD_i$) from CRT large–small and CRT above–below. M_i and $RTSD_i$ were computed after removing intra-individual outliers ($M_i \pm 4*RTSD_i$).

Results

SWAN Scores, Age, and Gender

For the general population sample, mean SWAN scores were $M \pm SD = 3.65 \pm 0.98$ for SWAN-AD, $M \pm SD = 3.75 \pm 1.00$ for SWAN-HI, and $M \pm SD = 3.70 \pm 0.92$ for SWAN-TOT. Mean scores on item level laid between 3.01 and 3.98 ($SD = 1.1$-1.4).

A univariate two-way ANOVA revealed no significant age, gender, or Age × Gender effects on SWAN-TOT in the general population sample, age: $F(1, 339) = 1.83$, n.s., $\eta_p^2 = .005$; gender: $F(1, 339) < 1$, n.s., $\eta_p^2 = .001$; Age ×

Gender: $F(1, 339) = 1.28$, n.s., $\eta_p^2 = .004$ (see Figure 1a). For SWAN-AD, gender effects were marginally significant, $F(1, 339) = 3.13$, $p = .08$, $\eta_p^2 = .009$, and a significant Age × Gender interaction was observed, $F(1, 339) = 6.25$, $p = .01$, $\eta_p^2 = .018$. While the correlation between age and SWAN-AD score was significant and positive for females ($r_{age, SWAN-AD} = .19$, $p < .01$), it was non-significant for male participants ($r_{age, SWAN-AD} = -.09$, n.s.; Figure 1b). Regarding SWAN-HI, only the age effect was marginally significant, $F(1, 339) = 3.46$, $p = .06$, $\eta_p^2 = .010$, indicating slightly higher levels of hyperactivity and impulsivity in younger than in older participants ($r_{age, SWAN-HI} = .11$, $p = .04$; Figure 1c).

Distribution

SWAN-TOT as well as SWAN-AD and SWAN-HI mean scores were approximately normally distributed as revealed by a visual inspection of histogram and quantile–quantile plots. Skew and kurtosis were small (SWAN-AD: skew = −0.13, kurtosis = 0.19; SWAN-HI: skew = 0.23, kurtosis = −0.30; SWAN-TOT: skew = 0.16, kurtosis = 0.003) and not significant (z scores $< \pm 1.96$).

Reliability

Cronbach's αs were .91, .94, and .95 for SWAN-AD, SWAN-HI, and SWAN-TOT respectively, indicating excellent internal consistency of the subscales and the total score. Every item was highly correlated with the entire (sub)scale (SWAN-AD: rs = .68-.82; SWAN-HI: rs = .74-.87; SWAN-TOT: rs = .60-.82). Test–retest reliability for a time interval of approximately 27 weeks ($M \pm SD = 26.95 \pm 2.39$; 22.29-32.29) indicated high stability for the total score and the subscales of SWAN (SWAN-AD: $r = .77$; SWAN-HI: $r = .79$; SWAN-TOT: $r = .81$).

Validity

A principal component analysis (PCA) revealed two components (eigenvalues > 1), which explained 65% of the variance. Item loadings and communalities are presented in Table 1. All SWAN-HI items loaded on Component 1 explaining 35% variance, and all SWAN-AD-items loaded on Component 2 explaining 30% variance. Cross loadings were small (<.30), except for Item 3.

In addition, the factorial structure of SWAN was examined in a CFA. The model revealed an acceptable fit, $\chi^2(117)$ = 306.27, $p < .001$, χ^2/df = 2.62; root mean square error of approximation (RMSEA) = .069 (90% confidence interval [CI] = [.059, .078]); comparative fit index (CFI) = .939; standardized root mean square residual (SRMR) = .041. An examination of the modification indices showed that the model fit can be further improved by allowing correlated measurement errors between two pairs of semantically strongly related items: Item 5: "Organize tasks and activities" and Item 7: "Keep track of things necessary for activities," as well as Item 10: "Sit still" and Item 11: "Stay seated." This modification resulted in a significant improvement in model fit, Satorra–Bentler scaled $\Delta\chi^2(2)$ = 49.68, $p < .001$. The final model showed a satisfactory to good fit regarding all fit indices, $\chi^2(115)$ = 275.85, $p < .001$, χ^2/df = 2.40; RMSEA = .064 (90% CI = [.054, .074]); CFI = .948; SRMR = .041. Table 1 presents the standardized loadings of the variables on the three factors and R^2 values. While all items loaded significantly on G-ADHD and all attention-related items loaded significantly on SP-AD, only five of the nine activity/impulsivity items showed significant loading on SP-HI. This indicates that SWAN-HI was less homogeneous as a subscale than SWAN-AD.

Convergent validity was assessed by computing Pearson's correlation coefficients between SWAN scores and CBCL-AP. Significant negative correlations were observed for all SWAN scales (SWAN-AD: $r = -.53$; SWAN-HI: $r = -.43$; SWAN-TOT: $r = -.52$). Visual inspection of the scatter plots revealed heteroscedasticity, indicating that SWAN allowed further differentiation within the participants with a score of 0 in CBCL-AP (Figure 2). In addition, for the ADHD sample, SWAN scores were significantly negatively correlated to FBB-ADHS scores (Table 2).

Regarding the *clinical utility* of SWAN, a two-way ANCOVA revealed significant differences between patients and healthy participants in SWAN-AD, patients: $M \pm SD$ = 2.22 ± 0.89, controls: $M \pm SD$ = 3.74 ± 0.92, $F(1, 380)$ = 135.25, $p < .001$, η_P^2 = .262; SWAN-HI, patients: $M \pm SD$ = 2.54 ± 0.80, controls: $M \pm SD$ = 3.83 ± 0.95, $F(1, 380)$ = 96.53, $p < .001$, η_P^2 = .203; and SWAN-TOT, patients: $M \pm SD$ = 2.38 ± 0.72, controls: $M \pm SD$ = 3.78 ± 0.86, $F(1, 380)$ = 137.46, $p < .001$, η_P^2 = .266, even after controlling for gender, age, and the Gender × Age interaction. Thus, patients with ADHD showed more difficulties in controlling their attention and impulsivity. In addition, ROC analysis revealed large area under the curve (AUC) values (SWAN-AD: 0.90; SWAN-HI: 0.87; SWAN-TOT: 0.91), indicating an excellent ability of SWAN to differentiate between participants with and without ADHD. Cutoff scores for optimal sensitivity and specificity were identified via visual inspection of the ROC curve: SWAN-AD: 2.94 for 84% sensitivity and 81% specificity; SWAN-HI: 3.06 for 84% sensitivity and 75% specificity; SWAN-TOT: 2.86 for 84% sensitivity and 88% specificity.

As a further estimate of *concurrent validity*, the correlations between SWAN scores and ISV as well as PS were investigated in a SEM approach. The model yielded an acceptable to good fit, $\chi^2(182)$ = 391.67, $p < .001$, χ^2/df = 2.15; RMSEA = .058 (90% CI = [.050, .066]); CFI = .947; SRMR = .041. Correlations between SWAN scores and ISV were significant for the G-ADHD and for SP-AD, but not for SP-HI (see Figure 3). PS was significantly correlated with SP-AD, but not with G-ADHD or SP-HI. Constraining the correlation coefficients between SP-AD and ISV to be the same as the correlations between SP-AD and PS resulted in a marginally significant decrease of model fit, Satorra–Bentler scaled $\Delta\chi^2(1)$ = 3.47, $p = .06$, indicating that SP-AD and ISV were stronger related than SP-AD and PS. Constraining the correlation coefficients between G-ADHD and SP-HI to be the same for ISV and PS did not yield significant changes in model fit.

Discussion

The aim of the present study was to introduce the German translation of the SWAN Scale and to report its psychometric qualities based on data from a large sample of school-aged children and adolescents with and without an ADHD diagnosis. In sum, the reported results support the excellent reliability and validity of SWAN-DE, which are comparable with the original version (Arnett et al., 2013) and with other translations (Chan et al., 2014; Lakes et al., 2012; Robaey et al., 2007; Table 3).

Similar to previous results (e.g., Arnett et al., 2013), the scores of SWAN and its subscales were approximately normally distributed, which supports the view of ADHD symptomatic as a dimensional construct continuously distributed across the general population and confirms SWAN as a scale capturing both poles of the ADHD continuum. In addition, SWAN-DE showed excellent internal consistency, comparable with previous studies (Lakes et al., 2012) and high test–retest reliability. Considering the comparatively long test–retest interval of 27 weeks, the fact that SWAN-DE showed higher test–retest indices than the original SWAN and its Spanish version with a 12-weeks test–retest interval (Lakes et al., 2012) provides strong support for the high temporal stability of individual differences in SWAN-DE scores.

Table 1. PCA and CFA Results for SWAN-DE.

		PCA			CFA			
Item		C 1	C 2	h²	G-ADHD	SP-AD	SP-HI	R^2
1.	Give close attention to detail and avoid careless mistakes	−.20	**.86**	.54	0.43 (0.06)***	0.52 (0.05)***	—	.45
2.	Sustain attention on tasks or play activities	.09	**.74**	.64	0.62 (0.04)***	0.48 (0.05)***	—	.61
3.	Listen when spoken to directly	**.37**	**.49**	.63	0.72 (0.03)***	0.32 (0.05)***	—	.61
4.	Follow through on instructions and finish school work/chores	.18	**.68**	.66	0.66 (0.04)***	0.45 (0.05)***	—	.64
5.	Organize tasks and activities	−.15	**.90**	.65	0.51 (0.05)***	0.54 (0.06)***	—	.55
6.	Engage in tasks that require sustained mental effort	−.12	**.90**	.68	0.52 (0.05)***	0.61 (0.05)***	—	.64
7.	Keep track of things necessary for activities	−.10	**.90**	.69	0.55 (0.05)***	0.55 (0.06)***	—	.60
8.	Ignore extraneous stimuli	.22	**.50**	.44	0.55 (0.05)***	0.33 (0.06)***	—	.41
9.	Remember daily activities	.18	**.62**	.57	0.62 (0.04)***	0.37 (0.05)***	—	.52
10.	Sit still	**.69**	.16	.66	0.76 (0.03)***	—	0.10 (0.07)	.59
11.	Stay seated	**.82**	.07	.76	0.83 (0.03)***	—	0.25 (0.06)***	.75
12.	Modulate motor activity	**.91**	−.06	.76	0.83 (0.05)***	—	0.43 (0.09)***	.87
13.	Play quietly	**.97**	−.14	.77	0.82 (0.04)***	—	0.34 (0.11)**	.78
14.	Settle down and rest	**.75**	.06	.63	0.74 (0.04)***	—	0.18 (0.09)	.57
15.	Modulate verbal activity	**.95**	−.23	.66	0.74 (0.04)***	—	0.03 (0.13)	.55
16.	Reflect on questions	**.67**	.19	.66	0.83 (0.03)***	—	−0.21 (0.10)*	.74
17.	Await turn	**.85**	−.03	.68	0.83 (0.02)***	—	−0.11 (0.10)	.70
18.	Enter into conversations and games	**.52**	.27	.54	0.74 (0.04)***	—	−0.23 (0.08)**	.60

Note. PCA = principal components analysis; CFA = confirmatory factor analysis; G-ADHD = general ADHD factor; SP-AD = specific attention deficit factor; SP-HI = specific hyperactivity/impulsivity factor; SWAN-DE = Strengths and Weaknesses of ADHS Symptoms and Normal Behavior, German Version.
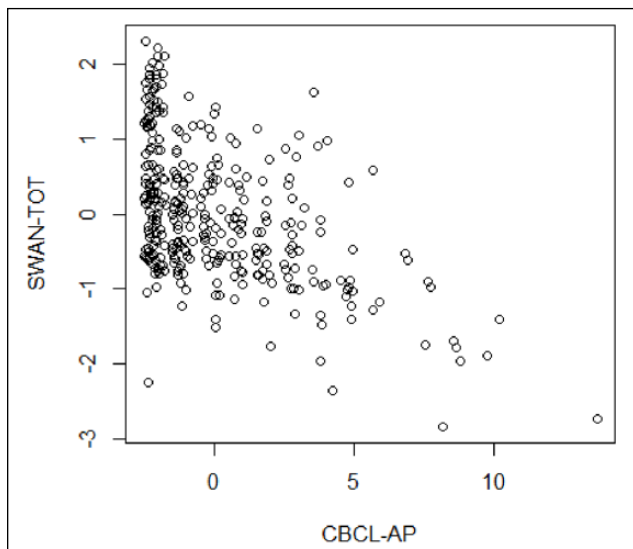

Figure 2. Correlation between SWAN-DE total score and CBCL-AP.
Note. CBCL-AP = Child Behavior Checklist, Attention Problems; SWAN-TOT = SWAN total score; SWAN = Strengths and Weaknesses of ADHD–Symptoms and Normal Behavior; SWAN-DE = Strengths and Weaknesses of ADHS Symptoms and Normal Behavior, German Version.

Alongside its excellent reliability, we were able to confirm the original structure of SWAN for its German translation

Table 2. Correlations Between the Sub-Scales of SWAN and FBB-ADHS.

	SWAN-AD	SWAN-HI	SWAN-TOT
FBB-ADHS-IN	−.67***	−.32***	−.59***
FBB-ADHS-HI	−.40***	−.67***	−.61***
FBB-ADHS-TOT	−.60***	−.56***	−.68***

Note. SWAN = Strengths and Weaknesses of ADHD–Symptoms and Normal Behavior; FBB = Fremdbeurteilungs-bogen für ADHS; SWAN-AD = SWAN subscale Attention Deficit; SWAN-HI = SWAN subscale Hyperactivity/impulsivity; SWAN-TOT = SWAN total score; IN = inattention; HI = hyperactivity/impulsivity; TOT = total score.

with PCA and CFA. The PCA revealed a (two-factor) simple structure with all items loading substantially on the expected factors and an amount of explained variance that was comparable with other studies using the parent version of SWAN (Lai et al., 2011). In addition, the amount of variance explained by a two-factor solution in SWAN (65%) exceeded PCA results based on other widely used scales (e.g., FBB-ADHS with 46%; Erhart, Döpfner, & Ravens-Sieberer, 2008). Likewise, the CFA confirmed the bifactor model of ADHD symptoms of the original SWAN (Normand et al., 2012).

Regarding the validity of SWAN, our results revealed significant associations between SWAN and widely applied clinical scales (CBCL-AP, FBB-ADHS) in the expected
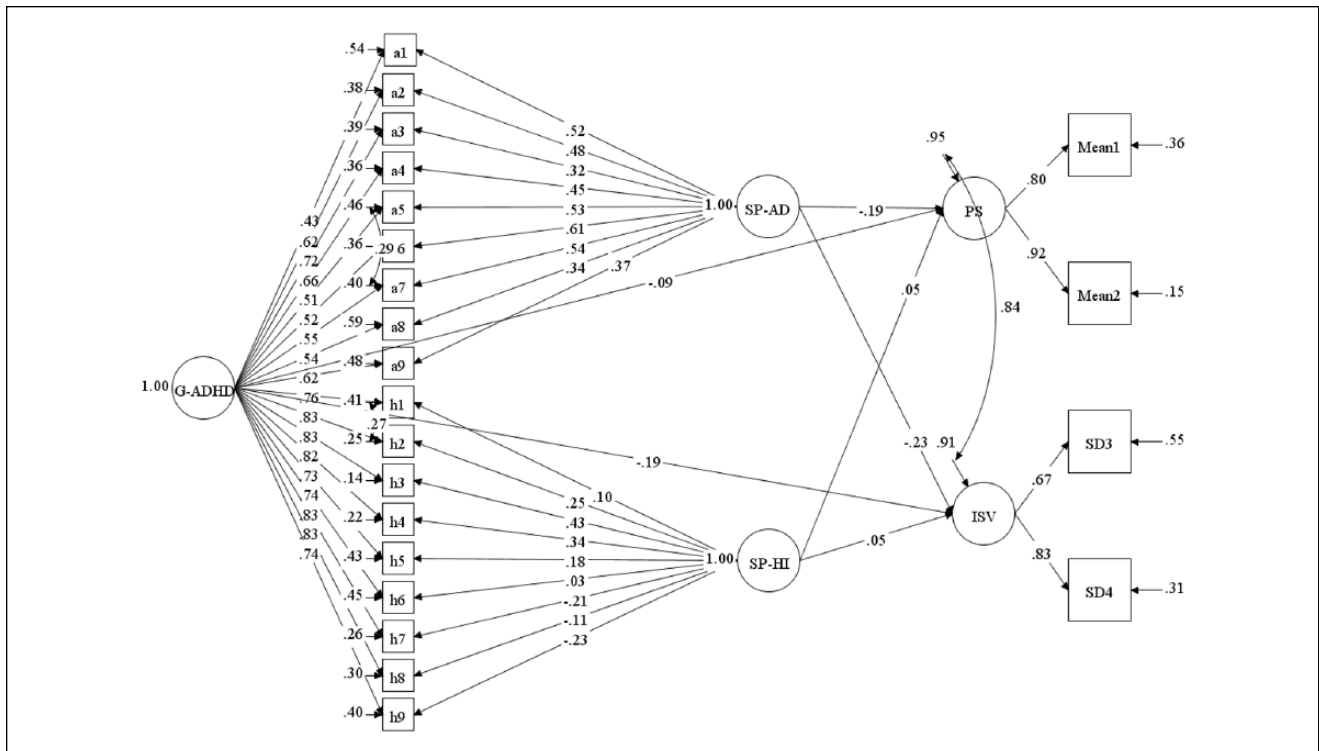
Figure 3. Structural model for the SWAN-DE.
Note. PS = processing speed; ISV = intra-subject variability; SP-AD = specific attention deficit factor; SP-HI = specific hyperactivity/impulsivity factor; G-ADHD = general ADHD factor; SWAN = Strengths and Weaknesses of ADHD–Symptoms and Normal Behavior; SWAN-DE = Strengths and Weaknesses of ADHS Symptoms and Normal Behavior, German Version.

direction. The magnitude of these correlations was comparable with previous studies based on the original scale and validated translations (Chan et al., 2014; Lakes et al., 2012; Robaey et al., 2007). In addition, we were able to replicate the observation that SWAN is able to differentiate between participants who receive mean scores of 0 on traditional clinical scales (Arnett et al., 2013; Polderman et al., 2007). These results underline SWAN's potential for population-based studies of ADHD, where discrimination in particular at the lower end of the distribution is required.

Regarding its clinical utility, SWAN-DE and particularly SWAN-TOT showed an excellent ability to differentiate between participants with and without ADHD. This observation is in accordance with previous studies, examining the French (Robaey et al., 2007) and the Chinese (Chan et al., 2014; Lai et al., 2011) versions of SWAN, revealing cutoffs with high diagnostic sensitivity and specificity. However, to provide clinically applicable cutoffs, a larger representative sample is needed. Nonetheless, the reported cutoffs illustrate the high potential of SWAN for clinical purposes.

In addition, the present study was able to confirm the relationship between ADHD traits assessed with SWAN and ISV, one of the key abnormalities and a candidate endophenotype of ADHD (Kuntsi & Klein, 2011), at the latent variable level in a general population sample. Thus,

increased ISV was associated with decreased ability to control ADHD behaviors in general and attention problems in particular. The observed magnitude of .23 for the correlation between ISV and SWAN-AD corresponds to Cohen's d of .47 (http://www.psychometrica.de) and lies below the 95% CI of [0.68, 0.84] for the effect size reported in a recent meta-analytic review (Kofler et al., 2013) for the group comparison between children with and without ADHD regarding ISV. To further investigate this result, we compared ISV and PS from the four CRTs between the patients with ADHD and a gender-, age-, and IQ-matched sample of healthy participants via linear mixed-effect models, including age as a covariate. These analyses revealed a significant effect of the diagnostic status regarding ISV, $b = -15.63$, $t(117) = -3.71, p < .001$, of substantial effect size ($r = .324$), corresponding to Cohen's d of .68, which lies within the reported interval by Kofler et al. (2013). Thus, the difference between the effect size reported by Kofler et al. (2013) and our results should rather be interpreted as a consequence of the correlational design of our study based on participants from the general population than as a result of specific task effects. Regarding PS, a similar, but overall weaker pattern of association with the specific attention problems factor was observed, which concurs with previous results (Klein, Wendling, Huettner, Ruder, & Peper, 2006).

Table 3. Overview of Validation Studies of the SWAN Scale.

| Study | Version | Distribution | Reliability | | Validity | | | |
			Cronbach's α	TRR	FV	CV	SP (%)	SE (%)
Swanson et al. (2006)	Original (TV)	normal	—	—	2 factors; 87%	—	—	—
Polderman et al. (2007)	Original (PV)	normal	—	—	—	−.38 to −.42 (CBCL-AP)	—	—
Robaey, Amre, Schachar, and Simard (2007)	French (PV)	normal	.88-.91	—	—	−.57 to −.68 (DISC-4) −.79 to −.85 (Conner's Scale PV) −.30 to −.32 (Conner's Scale TV)	88	86
Lakatos, Birkas, Toth, and Gervai (2010)	Hungarian (PV)	normal	.87-.93	—	—	.40 to .49 (CBCL-AP) .67 to .74 (SDQ-HI)	—	—
Lai et al. (2011)	Chinese (TV & PV)	normal	.90-.95 PV; .97-.98 TV	.84-.87 PV; .90-.92 TV	2 factors; 60% PV, 81% TV	—	66-89	55-83
Lakes, Swanson, and Riggs (2012)	Spanish (Sp) & Original (Or; PV)	normal	.91-.95 (Sp) .92-.95 (Or)	.49-.61 (Sp) .71-.76 (Or)	—	−.49 to −.54 (SDQ-HI)	—	—
Arnett et al. (2013)	Original (PV)	normal	.88	.57-.75	2 factors, 74%	.48 to .53 (DBRS)	98	58
Chan, Lai, Luk, Hung, and Leung (2014)	Chinese (TV & PV)	normal	—	—	—	.46 to .79 (DISC-4)	68-95	76-96
Present results	German (PV)	normal	.91-.95	.77-.81	2 factors, 65%	−.43 to −.53 (CBCL-AP) −.67 to −.68 (FBB-ADHD)	84	75-88

Note. SWAN = Strengths and Weaknesses of ADHD–Symptoms and Normal Behavior; TRR = test–retest reliability; FV = factorial validity; CV = convergent validity; SP = specificity; SE = sensitivity; TV = teacher version; PV = parent version; CBCL-AP = Child Behavior Checklist, subscale Attention Problems; DISC-4 = Diagnostic Interview Schedule for Children Version IV; SDQ-HI = Strengths and Difficulties Questionnaire, subscale Hyperactivity/Inattention; DBRS = Disruptive Behavior Rating Scale; FBB-ADHS = Fremdbeurteilungs-bogen für ADHS.

Thus, our results support the validity of SWAN-DE by revealing the associations between SWAN scores and ISV, and to a lesser degree PS, in the general population as described for the original SWAN (Crosbie et al., 2013).

Finally, several limitations of the present study should be discussed. Although the analyses were based on data from a wide-ranging sample regarding age and cognitive abilities, it was a preselected sample, consisting of healthy participants and posteriori supplemented by participant with ADHD. Thus, the sample may have not been representative for the general population of school-aged children in Germany, which could reduce the external validity of the results. However, comparing our results with the German norms for CBCL revealed a close correspondence: 96% of the girls and 94% of the boys younger than 12 years[1] as well as 97% of the girls and 98% of the boys older than 12 years in our sample received a CBCL-TOT score corresponding to a percentile rank (PR) of 98 or less. Thus, our sample was comparable with a representative sample of German children regarding cognitive, emotional, and behavioral problems. Second, although patients with ADHD were off medication for at least 24 hr before participating, receiving methylphenidate regularly may have affected the way their parents perceive the ADHD symptoms. In this case, however, patients would have been rated as less impaired and thus more similar to healthy controls, which would have reduced SWAN's ability to differentiate between participants with and without ADHD. Therefore, this limitation does not call into question the reported results, and further studies based on data from medication-naive participants could reveal an even higher clinical utility of SWAN.

Conclusion

The present study introduces SWAN-DE to the German-speaking community and confirms its excellent reliability and validity. SWAN-DE can now be used for research and clinical screening purposes.

Declaration of Conflicting Interests

Funding

Note

1. The fact that the percentage of young participants with a PR of 98 or less was slightly lower than expected could be

caused by differences in the age range between the norm sample (4-11 years) and our sample (8-11 years).

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 profile*. Burlington: University of Vermont.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Arnett, A. B., Pennington, B. F., Friend, A., Willcutt, E. G., Byrne, B., Samuelsson, S., & Olson, R. K. (2013). The SWAN captures variance at the negative and positive ends of the ADHD symptom dimension. *Journal of Attention Disorders*, *17*, 152-162.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*, 3186-3191.

Chan, G. F., Lai, K. Y., Luk, E. S., Hung, S.-F., & Leung, P. W.-L. (2014). Clinical utility of the Chinese Strengths and Weaknesses of ADHD-Symptoms and Normal-behaviors Questionnaire (SWAN) when compared with DISC-IV. *Neuropsychiatric Disease and Treatment*, *10*, 1533-1542.

Coghill, D., & Sonuga-Barke, E. J. S. (2012). Annual research review: Categories versus dimensions in the classification and conceptualisation of child and adolescent mental disorders—Implications of recent empirical study. *Journal of Child Psychology and Psychiatry*, *53*, 469-489.

Crosbie, J., Arnold, P., Paterson, A., Swanson, J., Dupuis, A., Li, X., . . . Schachar, R. J. (2013). Response inhibition and ADHD traits: Correlates and heritability in a community sample. *Journal of Abnormal Child Psychology*, *41*, 497-507.

Döpfner, M., Görtz-Dorten, A., & Lehmkuhl, G. (2008). *Diagnostik-System für Psychische Störungen im Kindes- und Jugendalter nach ICD-10 und DSM-IV, DISYPS-II*. Bern, Switzerland: Huber, Diagnostic system for psychic disorders in childhood and adolescence according to ICD-10 and DSM-IV, DISYPS-II.

Erhart, M., Döpfner, M., & Ravens-Sieberer, U. (2008). Psychometric properties of two ADHD questionnaires: Comparing the Conners' scale and the FBB-HKS in the general population of German children and adolescents—Results of the BELLA study. *European Child & Adolescent Psychiatry*, *17*(Suppl. 1), 106-115.

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*, 700-705.

Huss, M., Hölling, H., Kurth, B. M., & Schlack, R. (2008). How often are German children and adolescents diagnosed with ADHD? Prevalence based on the judgment of health care professionals: Results of the German Health and Examination Survey (KiGGS). *European Child & Adolescent Psychiatry*, *17*, 52-58.

Klein, C., Wendling, K., Huettner, P., Ruder, H., & Peper, M. (2006). Intra-subject variability in attention-deficit hyperactivity disorder. *Biological Psychiatry*, *60*, 1088-1097.

Kofler, M. J., Rapport, M. D., Sarver, D. E., Raiker, J. S., Orban, S., Friedman, L. M., & Kolomeyer, E. G. (2013). Reaction time variability in ADHD: A meta-analytic review of 319 studies. *Clinical Psychology Review*, *33*, 795-811.

Kuntsi, J., & Klein, C. (2011). Intraindividual variability in ADHD and its implications for research of causal links. In C. Stanford & R. Tannock (Eds.), *Behavioral neuroscience of attention deficit hyperactivity disorder and its treatment* (pp. 67-91). Springer-Verlag: Springer Publishers Berlin & Heidelberg.

Lai, K. Y. C., Leung, P. W. L., Luk, E. S. L., Wong, A. S. Y., Law, L. S. C., & Ho, K. K. Y. (2011). Validation of the Chinese strengths and weaknesses of ADHD-Symptoms and Normal-Behaviors Questionnaire in Hong Kong. *Journal of Attention Disorders*, *17*, 194-202.

Lakatos, K., Birkas, E., Toth, I., & Gervai, J. (2010). [Screening childhood behavior problems using short questionnaires II: The Hungarian version of the SWAN-scale (strengths and weaknesses of ADHD- symptoms and normal-behavior) for screening attention deficit/ hyperactivity disorder]. *Psychiatria Hungarica*, *25*, 493-502.

Lakes, K. D., Swanson, J. M., & Riggs, M. (2012). The reliability and validity of the English and Spanish strengths and weaknesses of ADHD and Normal Behavior Rating Scales in a preschool sample: Continuum measures of hyperactivity and inattention. *Journal of Attention Disorders*, *29*, 997-1003.

Levy, F., Hay, D. A., McStephen, M., Wood, C., & Waldman, I. (1997). Attention-deficit hyperactivity disorder: A category or a continuum? Genetic analysis of a large-scale twin study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*, 737-744.

Muthén, L. K., & Muthén, B. O. (2010). Mplus [Computer software]. Los Angeles, CA: Author.

Normand, S., Flora, D. B., Toplak, M. E., & Tannock, R. (2012). Evidence for a general ADHD factor from a longitudinal general school population study. *Journal of Abnormal Child Psychology*, *40*, 555-567.

Oberauer, K., Suß, H., Wilhelm, O., & Wittman, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. *Intelligence*, *31*, 167-193.

Polderman, T. J. C., Derks, E. M., Hudziak, J. J., Verhulst, F. C., Posthuma, D., & Boomsma, D. I. (2007). Across the continuum of attention skills: A twin study of the SWAN ADHD rating scale. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *48*, 1080-1087.

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Robaey, P., Amre, D., Schachar, R., & Simard, L. (2007). French version of the Strengths and Weaknesses of ADHD Symptoms and Normal Behaviors (SWAN-F) Questionnaire. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *16*, 80-89.

Swanson, J., Shuck, S., Mann, M., Carlson, C., Hartman, K., Sergeant, J., . . . McCleary, R. (2006). *Categorical and dimensional definitions and evaluations of symptoms of ADHD:*

The SNAP and SWAN Rating Scales. Irvine: University of California.

Author Biographies

Yoanna Schulz-Zhecheva is a PhD student at the University of Freiburg.

Manuel Voelkle is a professor for Psychological Research Methods at Humboldt University, Berlin.

André Beauducel is a professor for Psychological Methods, Assessment, and Evaluation at the University of Bonn.

Nadja Buch is a psychologist and worked as a masters student in the Department of Child and Adolescent Psychiatry, University of Freiburg.

Christian Fleischhaker is a professor in Child and Adolescent Psychiatry at the Department of Child and Adolescent Psychiatry, University of Freiburg.

Stephan Bender is a professor in Child and Adolescent Psychiatry and head of the Department of Child and Adolescent Psychiatry, University of Cologne.

Christopher W. N. Saville is a lecturer in Psychology in the School of Psychology, Bangor University.

Monica Biscaldi is a senior consultant in Child and Adolescent Psychiatry at the University of Freiburg.

Christoph Klein is a professor in Psychology at the University of Freiburg.