

Evidence for a General ADHD Factor from a Longitudinal General School Population Study

Sébastien Normand · David B. Flora ·
Maggie E. Toplak · Rosemary Tannock

Published online: 28 October 2011
© Springer Science+Business Media, LLC 2011

Abstract Recent factor analytic studies in Attention-Deficit/Hyperactivity Disorder (ADHD) have shown that hierarchical models provide a better fit of ADHD symptoms than correlated models. A hierarchical model includes a general ADHD factor and specific factors for inattention, and hyperactivity/impulsivity. The aim of this 12-month longitudinal study was to test the generalizability of the hierarchical models of ADHD within an elementary school population of 6–9 year old children (250 boys, 260 girls). Examination of differences as a function of informant (parent vs. teacher ratings), sex, and time was conducted. Six potential factor structures for the 18 items of the SWAN (Strengths and Weaknesses of ADHD-symptoms and Nor-

mal-behavior) scale were tested using confirmatory and exploratory factor analyses. Hierarchical models with a general ADHD factor and two or three specific factors best accounted for parent and teacher reports of symptoms for both boys and girls and at two time points separated by a 12-month interval. Findings indicate that the 18 SWAN items measure a common latent trait as well as orthogonal factors or dimensions of inattention and hyperactivity/impulsivity.

Keywords ADHD · Childhood · Inattention · Hyperactivity/impulsivity · General factor · Hierarchical model · School sample

Portions of this paper were presented at the 2011 International Society for Research on Child and Adolescent Psychology Biennial Conference, Chicago.

Electronic supplementary material The online version of this article (doi:10.1007/s10802-011-9584-5) contains supplementary material, which is available to authorized users.

S. Normand
Department of Psychology, Université du Québec en Outaouais,
Gatineau, Canada

D. B. Flora · M. E. Toplak
Department of Psychology, York University,
Toronto, Canada

R. Tannock (✉)
Neurosciences and Mental Health Research Program,
The Hospital for Sick Children,
555 University Avenue,
Toronto, Ontario M5G 1X8, Canada
e-mail: rosemary.tannock@utoronto.ca

R. Tannock
Ontario Institute for Studies in Education, University of Toronto,
Toronto, Canada

Attention-Deficit/Hyperactivity Disorder (ADHD) is characterized by developmentally inappropriate levels of inattention, impulsivity/hyperactivity, or both (American Psychiatric Association 2000). Individuals with ADHD represent a heterogeneous group who display substantial variation in the patterning and the severity of the 18 symptoms (Nigg et al. 2005). One approach to an optimal ADHD phenotype is to consider the unity and diversity of the symptoms of inattention and hyperactivity/impulsivity. Previous exploratory and confirmatory factor analytic studies have consistently found that inattention and hyperactivity/impulsivity are distinguishable but correlated dimensions in both clinical and population samples (Willcutt et al. *under review*). However, some researchers have found that this model is quite vulnerable to developmental changes in ADHD symptoms (Hart et al. 1995; Larsson et al. 2011), whereas others have shown that ADHD symptoms may be both quantitatively and qualitatively different depending on gender (Gershon 2002). There is also support for a correlated three-factor model of inattention, hyperactivity, and impul-

sivity (e.g., Proctor and Prevatt 2009). Nonetheless, current data overwhelmingly support the concurrent and predictive validity as well as clinical importance of the distinction between the symptom domains of inattention and hyperactivity/impulsivity in terms of developmental trajectory, comorbid mental disorders, academic and social impairment, and neuropsychological dysfunction (Willcutt et al. *under review*). In contrast with the perspective that inattention, hyperactivity, and impulsivity are fully separable dimensions, it is possible that most symptoms, currently clustered under two dimensions of Inattention and Hyperactivity/Impulsivity, co-occur and cluster under a general “ADHD” dimension. However, some symptoms from each dimension might cluster under one or more orthogonal factors. It is becoming evident that factor models of ADHD need to account for both concurrent overlap and separability between the ADHD symptom domains.

An alternative approach, which accounts for an overarching general ADHD factor as well as unique symptom dimensions, is a hierarchical factor model (also known as a “bifactor model”; Holzinger and Swineford 1937; Schmid and Leiman 1957) that has one general factor with additional specific factors. Some researchers posit that a single general ADHD factor accounts for covariation among all symptoms simultaneously with two (or three) separate, specific factors (i.e., inattention and hyperactivity/impulsivity; see Fig. 1) that vary

independently from the general factor (e.g., Toplak et al. 2009). The specific factors thus contribute independent covariance among the symptoms beyond that accounted for the general factor. Hence, the hierarchical model suggests that there are distinct, as well as overlapping, causal pathways that converge on ADHD. Support for this model may improve our understanding of ADHD as a heterogeneous condition and reconcile apparent contradictory findings of general (Sonuga-Barke et al. 2008) and separate (Sagvolden et al. 2005) influences on ADHD symptom domains.

Evidence of the superiority of the hierarchical model of ADHD over one-, two-, or three-factor non-hierarchical models has been demonstrated in general population samples of children, samples of children and adolescents with ADHD recruited from the community, and clinic-referred samples, using parent, teacher, adolescent, or clinician reports (e.g., Dumenci et al. 2004; Gibbins et al. 2011; Martel et al. 2010a, b; Toplak et al. 2009; A summary of studies on the hierarchical models of ADHD is presented in Suppl. Table 1). Despite their large representative samples, the two existing population-based studies have important limitations to note. First, Dumenci and colleagues (2004) only used teacher reports and were therefore not able to confirm their results with other key informants, such as parents. In contrast, Ullebø and colleagues (*under review*) tested only a hierarchical model with three specific factors (i.e., inattention and separate

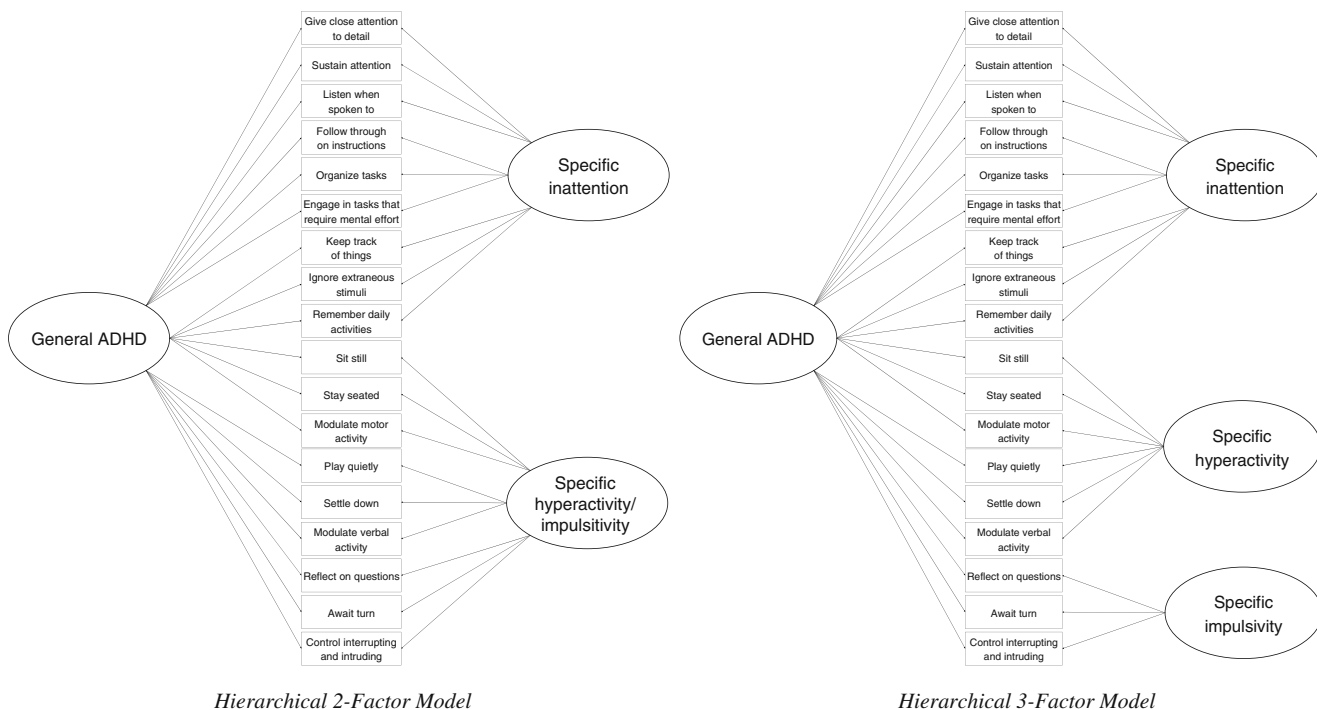


Fig. 1 Heuristic representations of the two hierarchical models of attention-deficit/hyperactivity disorder

Table 1 Descriptive statistics for the SWAN rating scale and SDQ in year 1: means and SDs (in parentheses)

	Parent (n=350)				Teacher (n=510)			
	Whole Sample	Boys	Girls	F or χ^2	Whole Sample	Boys	Girls	F or χ^2
SWAN								
Inattention	-0.46 (0.93)	-0.26 (0.93)	-0.66 (0.89)	16.87***	-0.12 (1.27)	0.29 (1.21)	-0.51 (1.21)	55.73***
Hyperactivity/Impulsivity	-0.51 (0.96)	-0.32 (0.94)	-0.69 (0.96)	13.28***	-0.27 (1.21)	0.13 (1.15)	-0.64 (1.14)	57.18***
General	-0.48 (0.88)	-0.29 (0.86)	-0.68 (0.85)	17.70***	-0.19 (1.18)	0.21 (1.10)	-0.58 (1.13)	63.10***
SDQ								
Total Difficulties	12.93 (3.29)	12.74 (3.36)	13.11 (3.22)	1.07	11.75 (2.82)	12.03 (3.04)	11.49 (2.57)	4.69*
Emotional Symptoms	1.84 (1.90)	1.76 (2.01)	1.91 (1.78)	0.53	1.21 (1.83)	1.23 (1.86)	1.19 (1.81)	0.05
Conduct Problems	2.49 (1.06)	2.40 (1.09)	2.58 (1.04)	2.58	2.20 (1.00)	2.29 (1.16)	2.12 (0.82)	3.46
Hyperactivity	3.94 (1.28)	3.93 (1.31)	3.96 (1.27)	0.06	4.25 (1.30)	4.47 (1.39)	4.03 (1.17)	14.87***
Peer Problems	4.66 (1.04)	4.66 (1.06)	4.66 (1.01)	0.00	4.09 (0.95)	4.04 (0.99)	4.14 (0.91)	1.37
Prosocial Behavior	8.70 (1.61)	8.40 (1.72)	9.00 (1.42)	12.51***	8.10 (2.16)	7.59 (2.30)	8.58 (1.88)	28.49***
Total Impact (% of children with an abnormal score) ^a	10.4	13.9	7.0	4.39*	11.0	16.4	5.8	14.74***

SWAN = Strengths and Weaknesses of ADHD symptoms and Normal-behavior. SDQ = Strengths and Difficulties Questionnaire. Negative means represent score above average on the SWAN, as impairment is rated as more positive in SWAN

^a Children were considered with abnormal scores if their Total Impact score was “at or above” the 90th percentile (Total Impact score ≥ 2)

hyperactivity and impulsivity factors). Furthermore, neither of these studies tested the stability of their models longitudinally. In light of the temporal instability of DSM-IV ADHD symptoms and subtypes and given that ADHD specific symptoms can wax and wane over time (Lahey et al. 2005; Larsson et al. 2011), it is also important to test the hierarchical models prospectively. Evidence of superiority of the hierarchical models in comparison to other models in longitudinal designs may confirm the well-accepted and empirically driven conceptualization that a latent general ADHD construct holds across time despite variation in the specific symptom presentation.

The unique objective of this study was to test the hierarchical models of ADHD in a normative sample of elementary school-aged children aged between 6 and 9, with a particular emphasis on potential informant, sex, and temporal differences over a 12-month period. To do so, we used a behavior rating scale designed to be sensitive to the full range of the behaviors of interest in a general school population: specifically the SWAN (Strengths and Weaknesses of ADHD-symptoms and Normal-behavior) scale (Swanson et al. 2001a). It is an empirical question whether hierarchical models fit data well in community samples as well as clinical ones. Evidence that they do would suggest that we are capturing an important structure or model of how these symptoms are organized across both general and clinical populations.

We therefore tested two hierarchical models with either two specific factors (inattention and hyperactiv-

ity/impulsivity) or three specific factors (inattention and separate hyperactivity and impulsivity factors). To test whether the hierarchical models are the optimal fitting models, we compared them to other factor analytic models published in the field of ADHD and examined them across informants. More specifically, we contrasted the fit of these hierarchical models with the fit of non-hierarchical models, namely one-factor model and confirmatory factor models with two and three correlated factors. Additionally, it is possible that confirmatory factor models in which each symptom is related to only one factor and its direct relationships with other factors (i.e., cross-loadings) are fixed to zero (as in our non-hierarchical correlated factor models) are likely overly restrictive for ADHD measurement scales. Instead, a non-hierarchical exploratory factor model in which each item is allowed to load on both an inattention and a hyperactivity factor may be more realistic (i.e., thus all cross-loadings are freely estimated; see, e.g., Asparouhov and Muthén 2009), in that some manifest symptoms might be influenced by both an inattention construct and a hyperactivity construct. Hence, with this non-hierarchical exploratory factor model each ADHD symptom is determined by not one but two factors, a characteristic it shares with the general-factor hierarchical model but not the strict non-hierarchical confirmatory factor model. Thus, the hierarchical model and the non-hierarchical exploratory model have a similar number of parameters and might be

expected to provide near-equivalent fit to data.¹ To investigate this possibility, we also compared the fit of the hierarchical models with that of a two-factor exploratory model. We predicted that, regardless of informant, sex, or time point, the hierarchical model, with one general factor and two specific factors, would be the best fitting model, and one which could account for both the overlap and separateness of the inattention, hyperactivity/impulsivity symptoms.

Method

Participants and Procedure

Participants were parents and teachers of 524 children (259 boys, 265 girls), aged 6 to 9 years ($M=7.60$; $SD=0.92$ in Year 1 and $M=8.61$; $SD=0.93$ in Year 2). The majority of children had English as their primary language (97%) and was primarily Caucasian (86%). Children were recruited from grades 1 to 3 in seven public elementary schools, which constitute 20% of the 33 schools in a large rural and suburban district school board in Southern Ontario, Canada. The data were collected as part of a 2-year prospective study investigating behavioral symptoms of attention, cognitive measures of attention, and academic outcomes in elementary school-aged children (SSHRC #410-2008-1052). Data collection for the present study occurred in November of each year to ensure that the children had sufficient time to settle down and for teachers to have observed and interacted with the children for 2–3 months.

Complete data for teacher ratings were available for 510 participants in Year 1 (i.e., 14 or 2.7% of the 524 consenting teachers did not complete the ratings) and 491 in Year 2 (i.e., 96% retention). Complete data for parent ratings were available for 350 participants in Year 1 (69% of sample) and 347 participants in Year 2 (68% of sample; 99% retention in Year 2). Whereas parents rated their children at both time points, different sets of teachers rated the same set of children in Year 1 and Year 2.

This study was approved by the University of Toronto, the Hospital for Sick Children, and the participating school boards. Principals interested in having their school participate contacted the research team to learn more about the study. Information sessions for teachers of grades 1 to 3 were provided along with study information and consent packages. Consenting teachers then received research

packages (containing a cover letter, study information, consent forms, and the study questionnaires) to give to parents. Eligibility criteria were: 1) education in mainstream classrooms in either English or French (25% were in French immersion); 2) no major sensory or physical impairment that would preclude a child from hearing the instructions or completing the assessment tasks; and 3) written informed consent from the child's teacher and parent and verbal assent from child.

Participating parents were mostly mothers (90%) and had diverse levels of education: less than a high school degree (2%), high school degree or equivalence (6%), some college (9%), college degree (32%), some university (11%), university degree (31%), or post-graduate college degree (9%). According to parental report, some children were reported to have the following learning and mental health issues: ADHD (4%), language impairment (3%), learning disability (3%), and a behavior problem (2%). We compared the profiles of consenting parents (and their children) who did versus those who did not complete the behavior ratings to test for any systematic differences. As indicated in the supplementary Table 2, participating parents were more likely to have English as their primary language and their children were more likely to be enrolled in French immersion classes. Children with non-participating parents were rated by their teacher as having more problems with inattention and hyperactivity (as measured on the SWAN), but fewer peer problems than children with non-participating parents. However, it is important to note that the teacher SWAN scores of children with or without completed parent questionnaires were both solidly in the average range (see Suppl. Table 2).

Among the 52 teachers who participated, 87% were females and 81% were Caucasian. Classroom size varied from 13 to 26 students with an average of 19.7 students per class ($SD=2.5$) in Year 1, whereas classroom size varied from 16 to 30 students with an average of 21.6 students per class ($SD=3.6$) in Year 2. The number of years of experience as a teacher ranged from 1 year to 33 years with an average of 14.6 years ($SD=7.9$). Also, 40% of teachers possessed additional qualifications in special education. All had taught at least one student with frequent problems in paying attention and concentrating, with about only 2% of teachers reported currently having no student with such problems. Moreover, 43% of teachers reported currently having several inattentive students in their classroom (1–3 students), 35% had quite a few (4–5 students), and 20% had many students with such problems (more than 5 students). As indicated in Table 1, and as expected from previous population studies, parents in our sample rated boys as having more ADHD symptoms than girls according to the SWAN. Teachers rated boys as having more ADHD symptoms and other problematic behavior

¹ A limitation of the exploratory factor model is that because of "rotational indeterminacy" (see, e.g., Fabrigar et al. 1999, p. 281), a two-factor model with a relatively high interfactor correlation and potentially very small cross-loadings cannot be mathematically distinguished from a rotation of the same model that produces a smaller interfactor correlation along with larger cross-loadings.

compared to girls, on both the SWAN and Strengths and Difficulties Questionnaire (SDQ). Both parents and teachers rated boys as having less prosocial skills than girls on the SDQ. Both informants rated at least twice more boys than girls as being severely impaired. However, when considering the whole sample, only 10–11% of the participating children were rated as being severely impaired, confirming that this sample is generally unimpaired in daily functioning (see Table 1).

Measures

SWAN Questionnaire The Strengths and Weaknesses of ADHD-symptoms and Normal-behavior Scale (SWAN; Swanson et al. 2001a) is a revised version of the SNAP-IV (Swanson 1995), a behavior rating scale measuring inattentive, hyperactive, and impulsive behaviors as specified in DSM-IV-TR (American Psychiatric Association 2000). The SWAN is also based on the 18 ADHD symptoms listed in the DSM-IV-TR, divided into two subscales of nine items corresponding to the domains of Inattention (items 1–9) and Hyperactivity/Impulsivity (items 10–18). The SWAN differs from most behavior rating scales used for assessing developmental psychopathology in that the symptoms of ADHD are reworded to reflect relative strengths (better than average) as well as weaknesses (worse than average). For instance, the DSM-IV symptom “Often avoids, dislikes, or reluctantly engages in tasks requiring sustained mental effort” is reworded as “Engage in tasks that require sustained mental effort”; “Often has difficulty awaiting turn” is rephrased as “Await turn (stand in line and take turns)”. The items are rated on a seven-point scale (–3 = far above average; –2 = above average; –1 = somewhat above average; 0 = average; 1 = somewhat below average; 2 = below average; 3 = far below average; Swanson et al. 2001a). Thus, in line with the conventions for positive scores to reflect weaknesses or problem behaviors, positive scores on the SWAN indicate report of worse than average ADHD symptoms, whereas negative scores indicate better than average behavior. Children’s total scores ranged from a minimum of –27 to a maximum of 27 for each subscale. A significant advantage of the SWAN is that it yields an approximate normal distribution of the subscale total scores (Cornish et al. 2008; Hay et al. 2007; Polderman et al. 2007; Robaey et al. 2007; Swanson et al. 2001a) and avoids potential psychometric flaws that are associated with skewed distributions such as over-identifying extreme cases of ADHD (Hay et al. 2007; Swanson et al. 2001b). ADHD represents one tail of the normal distribution of ADHD symptoms (Levy et al. 1997), but it is highly likely that some children in the general population will have better than average self-regulation of attention, activity and

impulse control, and thus fall at the opposite tail of the distribution (Swanson et al. 2001a).

In our sample, the total scores (i.e., mean of items) from the SWAN in Year 1 were approximately normal for the overall General ADHD scale (Parent version: skewness=–0.08, kurtosis=0.22; Teacher version: skewness=–0.23, kurtosis=–0.07) and the Inattention (Parent version: skewness=0.19, kurtosis=0.53; Teacher version: skewness=–0.09, kurtosis=–0.24) and Hyperactivity/Impulsivity (Parent version: skewness=–0.22, kurtosis=0.10; Teacher version: skewness=–0.14, kurtosis=0.16) subscales.² The internal consistency coefficients in Year 1 for the General ADHD scale and the Inattention and Hyperactivity/Impulsivity subscales were homogeneously excellent, ranging from 0.93 to 0.98 for parents’ and teachers’ ratings. Pearson correlation coefficients between the parent SDQ hyperactivity scale (which includes 5 items that probe for symptoms of inattention and hyperactivity/impulsivity) and the SWAN Inattention, Hyperactivity, and General scales were 0.28, 0.39, and 0.36, respectively in Year 1. Similarly, corresponding correlations between the teacher SDQ hyperactivity scale and the SWAN Inattention, Hyperactivity, and General scales were 0.30, 0.36, and 0.35, respectively in Year 1. All correlations were significant (in all cases $p < 0.01$). Virtually identical results were found with the Year 2 data in terms of the SWAN scales’ skewness, kurtosis, internal consistency, and correlations with the SDQ hyperactivity subscale.

These psychometric properties are consistent with those of previous studies (e.g., Robaey et al. 2007; So et al. 2008). Good concurrent validity exists between the SWAN and other clinical tools that use the exact wording of the DSM-IV criteria for ADHD. For instance, different structured clinical interviews (e.g., Schedule for Affective Disorders and Schizophrenia for School-Age Children—Present and Lifetime Version—KSADS-PL, Kaufman et al. 1997; Diagnostic Interview Schedule for Children 4.0, Shaffer et al. 2000) and symptom-based parent and teacher ratings (e.g., Conners’ Rating Scales revised, Conners et al. 1998a, b) has been previously reported in the literature (e.g., Cornish et al. 2005; Robaey et al. 2007; Toplak et al. 2009). For example, correlations between the Conners teacher Inattention and Hyperactivity/Impulsivity subscales and SWAN subscales were strong and ranged from 0.87 to 0.91 whereas they ranged from 0.75 to 0.82 on the parent version of the scales (Cornish et al. 2005).

Strengths and Difficulties Questionnaire We used the Strengths and Difficulties Questionnaire (SDQ; Goodman

² Here we report skewness and kurtosis for total Inattention and Hyperactivity/Impulsivity scores for descriptive purposes. Skewness and kurtosis of individual items, as relevant to model-fitting analyses, are reported in the Data Analysis section.

1997) as a screening measure of social, emotional and behavioral functioning, as well as of impairment. Each of the 25 items of the SDQ has the response options of 0 (not at all), 1 (a little, sometimes) or 2 (very much, all of the time). Responses for each item are grouped into one of five 5-item subscales (prosocial, hyperactivity, emotional, conduct and peer problems), and a total difficulties score is also obtained. Besides covering common areas of emotional and behavioral difficulties, it also enquires whether the informant thinks that the child has a problem in these areas and, if so, asks about resultant distress and impairment at home and school. Each impairment question is rated from 0 (not at all/only a little), to 1 (quite a lot), to 2 (a great deal). The maximum score for parent-rated impairment scale is 10 (5 questions) whereas the maximum score for teacher-rated impairment is 6 (3 questions). The SDQ is well established in research on the mental health of children and adolescents and there is good evidence of its psychometric properties across different cultures. Results from a recent review of 48 studies ($N=131,223$) on reliability and validity of the parent and teacher SDQ confirmed strong psychometric properties, including adequate internal consistency, test–retest reliability, inter-rater agreement, construct validity, concurrent validity, and discriminant validity for the parent and teacher versions (Stone et al. 2010).

In our study, we used the SDQ Hyperactivity Scale to estimate the number of children in our sample in Year 1 who would probably be at-risk for ADHD classification; its internal consistency was good (parent version, Cronbach's $\alpha=0.84$; teacher version, Cronbach's $\alpha=0.90$). More specifically, we used the following modified SDQ predictive algorithm: a) parent “or” teacher SDQ Hyperactivity Scale ≥ 6 (corresponding to the 90th percentile); b) some hyperactivity/inattention symptoms rated on the other informant (parent or teacher) as indicated by scores on the SDQ Hyperactivity Scale ≥ 4 ; and c) manifest impairment, as rated by both parent and teacher on SDQ impact scale (scores of 2 or more; representing the 90th percentile). This scoring algorithm takes both symptom severity and level of impairment into account, and uses both parent and teacher ratings to indicate problems in more than one context. According to this algorithm, 7 children (1%; 5 boys, 2 girls) would meet “probable ADHD” criteria in our general community sample.

Data Analysis

The 18 symptoms used to assess ADHD from the SWAN parent and teacher reports were used for the present analyses. Analyses were conducted separately for each informant on the questionnaire. All teacher-report models were fitted using aggregated analysis under complex

sampling as described by Muthén and Satorra (1995) to account for non-independence of observations due to each teacher rating multiple students in the class. Confirmatory factor analysis (CFA) was utilized to test six potential factor structures underlying the 18 ADHD symptoms: a) a one-factor model of inattention/hyperactivity/impulsivity; b) a non-hierarchical two-factor model with correlated inattention and hyperactivity/impulsivity factors; c) a non-hierarchical three-factor model with correlated inattention, hyperactivity, and impulsivity factors; d) a hierarchical model of a general ADHD factor with two specific factors of inattention and hyperactivity/impulsivity; e) a hierarchical model of a general ADHD factor with three specific factors of inattention, hyperactivity, and impulsivity, and f) a non-hierarchical exploratory two-factor model. As explained above, hierarchical models differ from non-hierarchical models in that they account for the covariation among all ADHD symptoms in terms of a general factor capturing the shared variation of all items, along with separate, uncorrelated specific factors capturing residual covariance among particular subgroups of items (see Fig. 1). Alternatively, non-hierarchical models include only factors for subgroups of items, and the relationships among items from different subgroups are accounted for by allowing these factors to correlate.

Because the SWAN items have a seven-point response scale, these items were factor analyzed as continuous variables with traditional maximum likelihood estimation because research shows this method is preferable with seven-category ordinal variables (Finney and DiStefano 2006). Although each of the individual symptoms was approximately normally distributed (across items, parent report univariate skewness ranged from -0.27 to 0.28 and univariate kurtosis ranged from -0.45 to 1.13 for both years while teacher report skewness ranged from -0.19 to 0.20 and kurtosis from -0.64 to 0.74), to account for potential multivariate non-normality, Satorra-Bentler fit statistics and robust standard errors (Satorra and Bentler 1988) were used. All models were estimated using Mplus (version 5.2; Muthén and Muthén 2002). Model fit was evaluated using the root mean square error of approximation (RMSEA), comparative-fit index (CFI), and Tucker-Lewis index (TLI) with acceptable model fit indicated by RMSEA values of 0.08 or lower along with CFI and TLI values of 0.95 or higher (Yu and Muthén 2002).

To test factorial invariance across sex, we followed procedures outlined by Widaman and Reise (1997) using a series of nested multiple-group CFA models. Although nested model comparisons often rely on χ^2 difference tests, recent methodological research suggests that examining alternative fit indices is preferable for this purpose because the χ^2 statistic is overly sensitive to sample size and ignores model parsimony. Thus, following Cheung and Rensvold

(2002) and Chen (2007) in the current analyses two models were considered to have equivalent fit if the decrease in CFI (Δ CFI) was 0.01 or less and if the increase in RMSEA (Δ RMSEA) was not greater than 0.01. First, to establish configural invariance, we examined the fit of a two-group model in which the basic model specification was identical for boys and girls, but all parameters were free to vary across sex (except those needed for overall model identification). Next, to test weak invariance, the fit of the initial configural invariance model was compared to that of a model with all factor loadings constrained to be equal across sex. If weak invariance held, we then tested strong invariance by comparing the fit of the weak invariance model to that of a model with all factor loading and all symptom intercept parameters constrained to equality across groups.

To test for longitudinal factorial invariance by year, we compared a series of single-group hierarchical models following Widaman et al. (2010). Specifically, a given model contained a set of general and specific factors for both Year 1 symptoms and Year 2 symptoms. The Year 1 general factor freely covaried with the Year 2 general factor, and likewise each Year 1 specific factor freely covaried with the corresponding Year 2 specific factor but all other interfactor covariances were constrained to zero. Additionally, the residual term for each Year 1 symptom was allowed to covary with that for the corresponding symptom from Year 2. Given this structure, we again fitted models

representing configural, weak, and strong invariance based on whether factor loading and intercept parameters for Year 1 symptoms were constrained to equal those for the corresponding Year 2 symptoms.

Results

Model Selection

Our results for parent- and teacher-report data (see Table 2) suggest that both the hierarchical 2-factor model and the hierarchical 3-factor model have better fit compared to the simple one-factor model and the non-hierarchical, correlated factor models. As detailed in Table 2, for Year 1 parent data, the hierarchical 2-factor model had slightly better fit (CFI=0.96, TLI=0.95, RMSEA=0.057) than the hierarchical 3-factor model (CFI=0.96, TLI=0.94, RMSEA=0.059), while the reverse occurred with Year 2 data (hierarchical 2-factor model: CFI=0.96, TLI=0.95, and RMSEA=0.064; hierarchical 3-factor model: CFI=0.97, TLI=0.96, and RMSEA=0.057). In Year 1 teacher data, the hierarchical 3-factor model best fit the data (hierarchical 2-factor model: CFI=0.94, TLI=0.93, and RMSEA=0.081; hierarchical 3-factor model: CFI=0.96, TLI=0.94, and RMSEA=0.072), but in Year 2 the hierarchical 2-factor model had the best fit (hierarchical 2-factor model:

Table 2 Fit of ADHD symptom CFA models for parent and teacher SWAN

Model	Parent report						Teacher report					
	df	χ^2	CFI	TLI	RMSEA	RMSEA 90% CI	df	χ^2	CFI	TLI	RMSEA	RMSEA 90% CI
Year 1 ^a												
One-factor model	135	892.69	0.77	0.74	0.127	(0.119, 0.135)	135	1998.04	0.74	0.70	0.164	(0.158, 0.171)
Correlated 2-factor	134	489.24	0.89	0.88	0.087	(0.079, 0.095)	134	838.16	0.90	0.89	0.102	(0.095, 0.108)
Correlated 3-factor	132	437.46	0.91	0.89	0.081	(0.073, 0.090)	132	642.76	0.93	0.92	0.087	(0.080, 0.094)
Hierarchical 2-factor	117	252.05	0.96	0.95	0.057	(0.048, 0.067)	117	508.48	0.94	0.93	0.081	(0.074, 0.088)
Hierarchical 3-factor	117	260.04	0.96	0.94	0.059	(0.049, 0.069)	117	425.63	0.96	0.94	0.072	(0.065, 0.079)
Exploratory 2-factor	118	395.81	0.92	0.89	0.082	(0.073, 0.091)	118	649.81	0.92	0.90	0.094	(0.087, 0.101)
Year 2 ^b												
One-factor model	135	903.01	0.81	0.79	0.128	(0.120, 0.136)	135	2077.43	0.76	0.72	0.171	(0.165, 0.178)
Correlated 2-factor	134	431.69	0.93	0.92	0.080	(0.072, 0.089)	134	845.11	0.91	0.90	0.104	(0.097, 0.111)
Correlated 3-factor	132	380.99	0.94	0.93	0.074	(0.065, 0.082)	132	588.37	0.94	0.93	0.084	(0.077, 0.091)
Hierarchical 2-factor	117	283.07	0.96	0.95	0.064	(0.054, 0.073)	117	368.68	0.97	0.96	0.066	(0.059, 0.074)
Hierarchical 3-factor*	118	252.74	0.97	0.96	0.057	(0.048, 0.067)	117	382.46	0.97	0.96	0.068	(0.061, 0.076)
Exploratory 2-factor	118	352.04	0.94	0.93	0.076	(0.067, 0.085)	118	627.47	0.94	0.92	0.094	(0.087, 0.101)

df degrees of freedom; χ^2 chi-square fit statistic; CFI Comparative Fit Index; TLI Tucker-Lewis Index; RMSEA Root Mean Square Error of Approximation; CI Confidence Interval

*For this model, the residual variance term for symptom 11 was constrained to zero to prevent the occurrence of an improper solution using the parent reports (i.e., Heywood case)

^a Year 1, parent: n=350; teacher n=510

^b Year 2, parent: n=347; teacher n=491

CFI=0.97, TLI=0.96, and RMSEA=0.066; hierarchical 3-factor model: CFI=0.97, TLI=0.96, and RMSEA=0.068). However, within both years, the fit of each hierarchical model for both parent and teacher SWAN was very similar. We therefore conclude that the hierarchical models with one general factor and either two or three specific factors (see Fig. 1) fit the data better than all competing models, using both parent- and teacher-report data in Year 1. This conclusion is cross-validated with the follow-up data from Year 2 (see Table 2). However, it is apparent that all of these models fit better for parents than for teachers.

Hierarchical Two-Factor Model Interpretation

With the parent-rating data, all symptoms had significant, positive loadings on the general ADHD factor, with completely standardized loadings ranging from 0.51 to 0.91 (all $ps < 0.05$; see Table 3 for standardized factor loading estimates and R^2 values). With the exception of a few symptoms from the inattention symptom domain (e.g., “listen when spoken to directly”, “ignore extraneous stimuli”, and “sustain attention on tasks or play activities”) that were more strongly related to the general factor than their specific factor, most inattention symptoms were

related similarly to the general factor and their specific factor. In contrast, all symptoms from the hyperactivity/impulsivity domain were more strongly related to the general factor than their specific factor. With the teacher-rating data, all symptoms had significant, positive loadings on the general ADHD factor, with completely standardized loadings ranging from 0.67 to 0.95 (all $ps < 0.05$; see Table 3 for standardized factor loading estimates and R^2 values). All symptoms from the inattention symptom domain were more strongly related to the general factor than their specific factor. Similarly, all symptoms from the hyperactivity/impulsivity domain were more strongly related to the general factor than their specific factor. The most important implication from these findings is that a general factor substantially influences all of the ADHD symptoms while the specific factors account for additional variance not from this general factor.

The interpretation of the standardized factor loading estimates for the hierarchical three-factor model in Year 1 was similar to those of the hierarchical two-factor model (see Table 4). Similarly, the overall pattern of factor loadings of the hierarchical models in Year 2 was virtually identical to our findings in Year 1 (see Suppl. Tables 3 and 4).

Table 3 Completely standardized factor loadings for hierarchical 2-factor model fitted to year 1 parent-and teacher-report data

Symptom	Factor						R^2	
	General		Specific Ina		Specific Hyp/Imp			
	P	T	P	T	P	T	P	T
1. Give close attention to details and avoid careless mistakes.	0.53	0.67	0.59	0.52			0.62	0.72
2. Sustain attention on tasks or play activities.	0.69	0.79	0.48	0.49			0.70	0.86
3. Listen when spoken to directly.	0.70	0.83	0.30	0.27			0.58	0.77
4. Follow through on instructions & finish school work/chores.	0.60	0.79	0.51	0.52			0.61	0.89
5. Organize tasks and activities.	0.57	0.78	0.63	0.51			0.71	0.87
6. Engage in tasks that require sustained mental effort.	0.59	0.77	0.56	0.55			0.67	0.89
7. Keep track of things necessary for activities.	0.54	0.79	0.56	0.49			0.61	0.86
8. Ignore extraneous stimuli.	0.62	0.83	0.36	0.34			0.51	0.81
9. Remember daily activities.	0.51	0.72	0.51	0.44			0.51	0.71
10. Sit still (control movement of hands/feet or control squirming).	0.86	0.95			-0.20	-0.05	0.78	0.91
11. Stay seated (when required by class rules/social conventions).	0.91	0.94			-0.30	-0.02	0.92	0.89
12. Modulate motor activity (inhibit inappropriate running/climbing).	0.76	0.91			-0.03	0.07	0.58	0.83
13. Play quietly (keep noise level reasonable).	0.71	0.89			0.22	0.23	0.55	0.85
14. Settle down and rest (control constant activity).	0.77	0.93			0.22	0.11	0.64	0.88
15. Modulate verbal activity (control excess talking).	0.73	0.86			0.41	0.32	0.71	0.85
16. Reflect on questions (control blurting out answers).	0.79	0.85			0.36	0.43	0.75	0.91
17. Await turn (stand in line and take turns).	0.79	0.85			0.23	0.43	0.67	0.92
18. Enter into conversations & games (control interrupting/intruding).	0.68	0.83			0.34	0.41	0.58	0.86

Ina Inattention; *Hyp* Hyperactivity; *Imp* Impulsivity; *P* Parent; *T* Teacher

All estimates are significant, $p < 0.05$, except those in italics

Table 4 Completely standardized factor loadings for hierarchical 3-factor model fitted to year 1 parent-and teacher-report data

Symptom	Factor								R^2	
	General		Specific Ina		Specific Hyp		Specific Imp			
	P	T	P	T	P	T	P	T	P	T
1. Give close attention to details and avoid careless mistakes.	0.51	0.65	0.60	0.54					0.63	0.72
2. Sustain attention on tasks or play activities.	0.68	0.77	0.49	0.52					0.70	0.86
3. Listen when spoken to directly.	0.70	0.82	0.31	0.31					0.58	0.77
4. Follow through on instructions & finish school work/chores.	0.60	0.77	0.50	0.54					0.61	0.89
5. Organize tasks and activities.	0.55	0.77	0.64	0.54					0.72	0.87
6. Engage in tasks that require sustained mental effort.	0.58	0.75	0.57	0.58					0.67	0.89
7. Keep track of things necessary for activities.	0.54	0.77	0.56	0.52					0.61	0.86
8. Ignore extraneous stimuli.	0.60	0.82	0.38	0.37					0.51	0.81
9. Remember daily activities.	0.51	0.70	0.51	0.47					0.51	0.71
10. Sit still (control movement of hands/feet or control squirming).	0.81	0.94			0.34	-0.19			0.77	0.92
11. Stay seated (when required by class rules/social conventions).	0.85	0.93			0.49	-0.19			0.96	0.90
12. Modulate motor activity (inhibit inappropriate running/climbing).	0.74	0.91			0.17	0.04			0.57	0.84
13. Play quietly (keep noise level reasonable).	0.78	0.93			-0.16	0.28			0.63	0.94
14. Settle down and rest (control constant activity).	0.82	0.95			-0.13	0.12			0.70	0.92
15. Modulate verbal activity (control excess talking).	0.79	0.90			-0.24	0.19			0.69	0.84
16. Reflect on questions (control blurting out answers).	0.79	0.88					0.38	0.36	0.77	0.90
17. Await turn (stand in line and take turns).	0.77	0.88					0.37	0.39	0.74	0.93
18. Enter into conversations & games (control interrupting/intruding).	0.69	0.85					0.31	0.39	0.57	0.88

Ina Inattention; *Hyp* Hyperactivity; *Imp* impulsivity; *P* parent; *T* teacher
 All estimates are significant, $p < 0.05$, except those in italics

Parent versus Teacher Reports

Comparison of the parent and teacher reports indicates that the symptom factor loadings on the general ADHD factor and R^2 values appeared stronger for teachers than parents on all the hierarchical models in both years. However, we found similar findings in terms of the general factor as the models for both informants produced moderate to strong loadings (i.e., >0.40) on this factor. Across hierarchical models, informants, and time points, inattention and hyperactivity/impulsivity (whether considered together or separately) symptoms were consistently more strongly related to the general factor than to their respective separate factors. One exception is that some parent-reported inattention symptoms in Year 1 were similarly related to the general factor and their specific factor. The loadings on the specific hyperactivity/impulsivity factors were often weak and sometimes non-significant or even negative. Overall, these findings suggest a single, dominant general latent factor underlying all 18 SWAN items.

Measurement Invariance of Sex, Informant, and Time

We examined factorial invariance by sex for both the hierarchical two-factor model and the hierarchical three-

factor model separately for parent and teacher report and separately for Year 1 and Year 2 data. We found that strong measurement invariance holds across sex for both the hierarchical two-factor model and the hierarchical three-factor model and across both parent- and teacher-report (see Tables 5 and 6). This finding was cross-validated with the Year 2 follow-up data. Thus, the measurement properties of the SWAN are equivalent for boys and girls with respect to relating the ADHD constructs implied by the hierarchical factor models to the 18 observed symptoms. We also found that strong measurement invariance holds across Year 1 and Year 2 data for both the hierarchical two-factor model and the hierarchical three-factor model and across both parent- and teacher-report (see Tables 5 and 6). Thus, the measurement properties of the SWAN are stable across a 12-month period with respect to relating the ADHD constructs implied by the hierarchical factor models to the 18 observed symptoms.³

³ In this study, we offered no hypothesis regarding possible age effects given the very limited age range (6–9 years old) of our sample participants. However, post hoc analyses dividing the sample into grades 1, 2 and 3 yielded similar findings and provided evidence of the superiority of the hierarchical models, despite a smaller sample size (results available from authors).

Table 5 Results of factorial invariance testing for hierarchical 2-factor model

	Factorial invariance model									
	Configural		Weak invariance				Strong invariance			
	CFI	RMSEA	CFI	Δ CFI	RMSEA	Δ RMSEA	CFI	Δ CFI	RMSEA	Δ RMSEA
By Gender										
Parent report										
Hierarchical 2-factor (year 1)	0.962	0.058	0.955	0.007	0.059	−0.001	0.954	0.001	0.058	0.001
Hierarchical 2-factor (year 2)	0.962	0.065	0.955	0.007	0.066	−0.001	0.953	0.002	0.065	0.001
Teacher report										
Hierarchical 2-factor (year 1)	0.937	0.091	0.933	0.004	0.088	0.003	0.929	0.004	0.088	0.000
Hierarchical 2-factor (year 2)	0.962	0.074	0.960	0.002	0.070	0.004	0.960	0.000	0.069	0.001
By Year										
Parent report										
Hierarchical 2-factor	0.962	0.038	0.961	0.001	0.037	0.001	0.960	−0.001	0.038	−0.001
Teacher report										
Hierarchical 2-factor	0.959	0.055	0.960	−0.001	0.053	0.002	0.960	0.000	0.052	0.001

CFI Comparative Fit Index; RMSEA Root Mean Square Error of Approximation

Discussion

In the current longitudinal general population-based study, we evaluated the relative fit of six competing potential factor structures to determine the best model for the relationships among inattention, hyperactivity, and impulsivity symptoms, as reported by parents and teachers for elementary school-aged boys and girls using the SWAN scale. Hierarchical models with a general factor and two or three specific factors had the best fit regardless of informant, sex, or time point. These findings indicate a general ADHD factor that accounts for covariation among all symptoms characteristic of ADHD, as well as distinct specific factors that explain additional relationships among symptoms beyond general ADHD-related psychopathology. Our findings therefore corroborate those of other researchers documenting the superiority of general factor models in ADHD, using various ADHD rating scales and clinical interviews (see Suppl. Table 1). Complementing the results of previous studies that mostly have been conducted using clinical samples, our study provides further evidence that hierarchical models of ADHD also fit data well in a community sample of children. This suggests that such models are capturing an important structure or model of how the ADHD symptoms are organized across samples and populations. The fact that we found good fit for both hierarchical models was not surprising and is in line with Toplak et al.'s (2009) findings for a clinical sample of adolescents with the teacher SWAN. Those authors interpreted the general factor model with two specific factors as it is more parsimonious than the model with three specific factors. It is also interesting to note that in the current study,

hyperactivity/impulsivity factor loadings in the two-factor solution appear more consistent across parents and teachers than in the three-factor solution. In other words, positive loadings for parents tend to parallel positive loadings for teachers. However, and perhaps more importantly, a dominant general factor is present regardless of whether there are two or three specific factors.

Notably, both parent and teacher reports converged and confirmed the superiority of the general factor models in contrast to other models. However, it is apparent that neither the general factor models nor the other non-hierarchical models fit as well for teachers as for parents. We propose several explanations for this difference. First, parents and teachers base their ratings on different contextual demands (i.e., home, school respectively). Second, parents and teachers also have different time frames of experience with a particular child, which may influence their ratings. Third, teachers also have access a different reference population (e.g., the whole class of students) than parents (e.g., siblings). Nonetheless, the fact that there is measurement invariance across the two time points for teachers, as well as for parents provides evidence of temporal stability of the general factor models across informants. Furthermore, on both parent and teacher reports, the inattention and hyperactivity/impulsivity (whether considered together or separately) symptoms were generally more strongly related to the general ADHD factor than to their respective separate factors. The factor loadings on the specific hyperactivity/impulsivity factor(s) were often even weak and sometimes non-significant or even negative. Taken together, these general results highlight an integral association between symptoms of inattention and

Table 6 Results of factorial invariance testing for hierarchical 3-factor model

	Factorial invariance model									
	Configural		Weak invariance				Strong invariance			
	CFI	RMSEA	CFI	ΔCFI	RMSEA	ΔRMSEA	CFI	ΔCFI	RMSEA	ΔRMSEA
By Gender										
Parent report										
Hierarchical 3-factor (year 1)	0.963	0.058	0.954	0.009	0.059	−0.001	0.953	0.001	0.059	0.000
Hierarchical 3-factor (year 2)	0.965	0.062	0.961	0.004	0.061	0.001	0.959	0.002	0.061	0.000
Teacher report										
Hierarchical 3-factor (year 1)	0.944	0.086	0.943	0.001	0.081	0.005	0.940	0.003	0.081	0.000
Hierarchical 3-factor (year 2)	0.961	0.075	0.960	0.001	0.071	0.004	0.959	0.001	0.070	0.001
By Year										
Parent report										
Hierarchical 3-factor	0.964	0.037	0.963	0.001	0.037	0.000	0.961	0.002	0.037	0.000
Teacher report										
Hierarchical 3-factor	0.962	0.053	0.961	0.001	0.052	0.001	0.961	0.000	0.051	0.001

CFI Comparative Fit Index; RMSEA Root Mean Square Error of Approximation

hyperactivity/impulsivity, suggesting overlap among associated etiological pathways (Sonuga-Barke et al. 2008).

At a theoretical level, our findings converge with findings from both clinical and population studies which suggest a single, general latent construct of ADHD psychopathology underlying all 18 DSM-IV ADHD symptoms. Whereas these specific and independent dimensions of symptoms do exist, our findings imply that individual symptoms in each ADHD domain seem to more often co-occur rather than occur in isolation. This result can perhaps partly explain some of the inter-individual heterogeneity manifest in ADHD and is in line with research findings documenting important variability in neuropsychological endophenotypes of ADHD (Sonuga-Barke et al. 2010). This finding may also help explain why co-occurrence of impairments across these dimensions (as opposed to isolated deficits) is the norm (e.g., Conners 2008). Notably, the non-hierarchical exploratory factor model also permits overlap between inattention and hyperactivity/impulsivity symptoms by allowing cross-loadings; this is theoretically consistent with the notion of the general factor in the current hierarchical models tested.

Limitations and Future Directions

Ours is one of the three investigations on the hierarchical model of ADHD using a general population sample (Dumenci et al. 2004; Ullebø et al. under review) but among these, it is the only one adopting a longitudinal design. Nonetheless, there are three key limitations of the present study. First, our findings are limited to a specific age range (6 to 9) and so may not be generalizable to other age groups. However,

evidence of the superiority of the hierarchical models of ADHD has been established with diverse age groups, including children, adolescents, and adults (see Suppl. Table 1). Second, although our is the only known longitudinal study on the hierarchical models of DSM-IV-TR ADHD symptoms, it is important to keep in mind that we only had two time points that were separated by only 1 year. One year might not be enough time for developmental differences to emerge. Nonetheless, our results indicate temporal stability of the ADHD hierarchical models over a 1-year follow-up, despite the change in one of the informants (teachers). Finally, in the current study we examined the ADHD symptoms alone in order to better understand the symptoms as an entity in DSM-IV and because our sample was constituted of children from the general population.

It will be important in future studies to explore the factor structure of ADHD with that of other disruptive behavior disorders and other correlates of ADHD, including cognitive functions, motivational processes, and personality trait profiles (e.g., Martel et al. 2011). Regarding the hierarchical models in particular, it will be crucial to examine the criterion-related and incremental validity of the separate general and specific factors. For example, evaluating the hierarchical models incorporating ADHD and ODD in relation to other models is likely to improve understanding of the structural relationship between these two conditions and to have important implications for diagnostic assessment, prognosis, and treatment (Martel et al. 2010a). This kind of research will allow us to examine further consistencies and inconsistencies across clinical and community samples, to eventually identify where important differences are.

Clinical Implications

Notwithstanding the limitations indicated above, our study has important clinical implications for finding an optimal diagnostic system that would enable clinicians and researchers to recognize and describe meaningful heterogeneity in ADHD in the DSM-V. Our findings are consistent with the notion that children with ADHD are characterized by substantial interindividual heterogeneity (Sonuga-Barke 2002; Sonuga-Barke et al. 2010) with some children exhibiting high inattentive and hyperactive-impulsive symptoms and other children exhibiting a different kind of symptom profile with high symptoms in only one domain. But in line with a recent continuous alternative to nominal subtypes of ADHD (Lahey and Willcutt 2010), it seems critical that the DSM-V committees consider the general ADHD factor and the relative contributions of inattention and hyperactivity/impulsivity symptoms in the review of diagnostic criteria.

Acknowledgement This research was financially supported by a grant from the Social Sciences and Humanities Research Council of Canada (SSHRC #410-2008-1052) and the Canada Research Chair Program (RT). We express appreciation to the families and schools from the Peterborough Northumberland Clarington Catholic School Board and the Kawartha Pine Ridge District School Board who participated in our study. The dedicated assistance of Min-Na Hockenberry, Marisa Catapang, Peter Chaban, Sarah Anne Gray and Danielle Pigon is also gratefully acknowledged.

Disclosure Dr. Tannock has been a consultant for Eli Lilly, McNeil, Purdue, and Shire, for which she has received honoraria that are donated to The Hospital for Sick Children's Foundation to support ADHD research. Drs. Normand, Flora and Toplak report no financial relationships with commercial interests.

References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders DSM-IV-TR Fourth Edition* (4th ed.). American Psychiatric Publishing, Inc.
- Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*, 397–438.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*, 464–504.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255.
- Conners, C. (2008). *Conners 3rd edition: Manual*. Toronto: Multi-Health Systems Inc.
- Conners, C. K., Sitarenios, G., Parker, J. D., & Epstein, J. N. (1998a). Revision and restandardization of the Conners Teacher Rating Scale (CTRS-R): factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology*, *26*, 279–291.
- Conners, C. K., Sitarenios, G., Parker, J. D., & Epstein, J. N. (1998b). The revised Conners' Parent Rating Scale (CPRS-R): factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology*, *26*, 257–268.
- Cornish, K. M., Manly, T., Savage, R., Swanson, J., Morisano, D., Butler, N., et al. (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*, 686–698.
- Cornish, K. M., Wilding, J. M., & Hollis, C. (2008). Visual search performance in children rated as good or poor attenders: the differential impact of DAT1 genotype, IQ, and chronological age. *Neuropsychology*, *22*, 217–225.
- Dumenci, L., McConaghy, S. H., & Achenbach, T. M. (2004). A hierarchical three-factor model of inattention-hyperactivity-impulsivity derived from the attention problems syndrome of the teacher's report form. *School Psychology Review*, *33*, 287–301.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*, 272–299.
- Finney, S. J., & DiStefano, C. (2006). Nonnormal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269–314). Charlotte, NC: Information Age Publishing.
- Gershon, J. (2002). A meta-analytic review of gender differences in ADHD. *Journal of Attention Disorders*, *5*, 143–154.
- Gibbins, C., Toplak, M. E., Flora, D. B., Weiss, M. D., & Tannock, R. (2011). Evidence for a general factor model of ADHD in adults. *Journal of Attention Disorders*. doi:10.1177/1087054711416310.
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry*, *38*, 581–586.
- Hart, E. L., Lahey, B. B., Loeber, R., Applegate, B., & Frick, P. J. (1995). Developmental change in attention-deficit hyperactivity disorder in boys: a four-year longitudinal study. *Journal of Abnormal Child Psychology*, *23*, 729–749.
- 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.
- Holzinger, K. J., & Swineford, F. (1937). The Bi-factor method. *Psychometrika*, *2*, 41–54.
- Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., et al. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*, 980–988.
- Lahey, B. B., & Willcutt, E. G. (2010). Predictive validity of a continuous alternative to nominal subtypes of attention-deficit/hyperactivity disorder for DSM-V. *Journal of Clinical Child and Adolescent Psychology*, *39*, 761–775.
- Lahey, B. B., Pelham, W. E., Loney, J., Lee, S. S., & Willcutt, E. (2005). Instability of the DSM-IV Subtypes of ADHD from preschool through elementary school. *Archives of General Psychiatry*, *62*, 896–902.
- Larsson, H., Dilshad, R., Lichtenstein, P., & Barker, E. D. (2011). Developmental trajectories of DSM-IV symptoms of attention-deficit/hyperactivity disorder: genetic effects, family risk and associated psychopathology. *Journal of Child Psychology and Psychiatry*, *52*, 954–963.
- 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.
- Martel, M. M., Gremillion, M., Roberts, B., von Eye, A., & Nigg, J. T. (2010a). The structure of childhood disruptive behaviors. *Psychological Assessment*, *22*, 816–826.

- Martel, M. M., Von Eye, A., & Nigg, J. T. (2010b). Revisiting the latent structure of ADHD: is there a 'g' factor? *Journal of Child Psychology and Psychiatry*, *51*, 905–914.
- Martel, M. M., Roberts, B., Gremillion, M., Von Eye, A., & Nigg, J. T. (2011). External validation of bifactor model of ADHD: Explaining heterogeneity in psychiatric comorbidity, cognitive control, and personality trait profiles within DSM-IV ADHD. *Journal of Abnormal Child Psychology*. doi:10.1007/s10802-011-9538-y
- Muthén, L. K., & Muthén, B. O. (2002). *Mplus user's guide* (5th ed.). Los Angeles: Muthén and Muthén.
- Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. In P. Marsden (Ed.), *Sociological methodology* (pp. 216–316). Cambridge: Blackwell.
- Nigg, J. T., Willcutt, E. G., Doyle, A. E., & Sonuga-Barke, E. J. S. (2005). Causal heterogeneity in attention-deficit/hyperactivity disorder: do we need neuropsychologically impaired subtypes? *Biological Psychiatry*, *57*, 1224–1230.
- 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*, *48*, 1080–1087.
- Proctor, B. E., & Prevatt, F. (2009). Confirming the factor structure of attention-deficit/hyperactivity disorder symptoms in college students using student and parent data. *Journal of Learning Disabilities*, *42*, 250–259.
- 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.
- Sagvolden, T., Johansen, E. B., Aase, H., & Russell, V. A. (2005). A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD) predominantly hyperactive/impulsive and combined subtypes. *The Behavioral and Brain Sciences*, *28*, 397–419. discussion 419–468.
- Satorra, A., & Bentler, P. M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. *ASA Proceedings of the Business and Economic Section*, 308–313.
- Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, *22*, 53–61.
- 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.
- So, C. Y. C., Leung, P. W. L., & Hung, S.-F. (2008). Treatment effectiveness of combined medication/behavioural treatment with chinese ADHD children in routine practice. *Behaviour Research and Therapy*, *46*, 983–992.
- Sonuga-Barke, E. (2002). Psychological heterogeneity in AD/HD—a dual pathway model of behaviour and cognition. *Behavioural Brain Research*, *130*, 29–36.
- Sonuga-Barke, E. J. S., Sergeant, J. A., Nigg, J., & Willcutt, E. (2008). Executive dysfunction and delay aversion in attention deficit hyperactivity disorder: nosologic and diagnostic implications. *Child and Adolescent Psychiatric Clinics of North America*, *17*, 367–384.
- Sonuga-Barke, E., Bitsakou, P., & Thompson, M. (2010). Beyond the dual pathway model: evidence for the dissociation of timing, inhibitory, and delay-related impairments in attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*, 345–355.
- Stone, L. L., Otten, R., Engels, R. C. M. E., Vermulst, A. A., & Janssens, J. M. A. M. (2010). Psychometric properties of the parent and teacher versions of the Strengths and Difficulties Questionnaire for 4- to 12-year-olds: a review. *Clinical Child and Family Psychology Review*, *13*, 254–274.
- Swanson, J. M. (1995). *SNAP-IV scale*. Irvine: University of California Child Development Center.
- Swanson, J. M., Schuck, S., Mann, M., Carlson, C., Hartman, K., Sergeant, J. A., et al. (2001a). Over-identification of extreme behavior in the evaluation and diagnosis of ADHD/HKD. Available at: <http://www.adhd.net>. Accessed April 6, 2011.
- Swanson, J. M., Schuck, S., Mann, M., Carlson, C., Hartman, K., Sergeant, J. A., et al. (2001b). Categorical and dimensional definitions and evaluations of symptoms of ADHD: the SNAP and SWAN ratings scales. Available at: <http://www.adhd.net>. Accessed April 6, 2011.
- Toplak, M. E., Pitch, A., Flora, D. B., Iwenofu, L., Ghelani, K., Jain, U., et al. (2009). The unity and diversity of inattention and hyperactivity/impulsivity in ADHD: evidence for a general factor with separable dimensions. *Journal of Abnormal Child Psychology*, *37*, 1137–1150.
- Ullebø, A., Breivik, K., Gillberg, C., Lundervold, A., & Posserud, M. (under review). The dimensionality of ADHD in a total population of Norwegian school children.
- Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–325). Washington, DC: American Psychological Association.
- Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives*, *4*, 10–18.
- Willcutt, E. G., Nigg, J. T., Pennington, B. F., Solanto, M. V., Rohde, L. A., Tannock, R., et al. (under review). Validity of DSM-IV attention-deficit/hyperactivity disorder symptom dimensions and subtypes.
- Yu, C., & Muthén, B. (2002). Evaluation of model fit indices for latent variable models with categorical and continuous outcomes. Presented at the Annual meeting of the American Educational Research Association, New Orleans