Published in final edited form as:

Assessment. 2009 June ; 16(2): 193–208. doi:10.1177/1073191108326924.

Psychometric Properties of Teacher SKAMP Ratings From a Community Sample

Desiree W. Murray,
Duke University Medical Center

Regina Bussing, and
University of Florida

Melanie Fernandez
University of Florida, currently at New York University Medical Center

Wei Hou and **Cynthia Wilson Garvan**
University of Florida

James M. Swanson and
University of California, Irvine

Sheila M. Eyberg
University of Florida

Abstract

This study examines the basic psychometric properties of the Swanson, Kotkin, Agler, M-Flynn, and Pelham Scale (SKAMP), a measure intended to assess functional impairment related to attention deficit hyperactivity disorder, in a sample of 1,205 elementary students. Reliability, factor structure, and convergent, discriminant and predictive validity are evaluated. Results provide support for two separate but related subscales, Attention and Department, and provide evidence that the SKAMP predicts school functioning above and beyond symptoms alone. Boys, African American children, and children living in poverty are rated as having higher impairment scores than girls, Caucasian children, and more advantaged peers. Norm-referenced data are provided by gender, race, and parental concern level. This study supports the reliability and validity of the SKAMP in a large, diverse community sample and broadens its clinical utility.

Keywords

ADHD; teacher ratings; impairment; psychometrics

A number of rating scales have been developed and validated for assessing symptoms of attention-deficit/hyperactivity disorder (ADHD) that are helpful in informing diagnostic decision making and detecting behavioral changes over time (Collett, Ohan, & Myers, 2003). However, increased attention is being directed at the assessment of impairment, reflecting how a child functions across different domains of day-to-day activities. Although symptoms and impairment overlap, these constructs may be distinguished conceptually in that symptoms and their severity are thought to describe disorders, whereas functional impairment represents a

state of the individual and how he or she functions across different roles or settings (Bird, 1999). Measuring functional difficulties associated with ADHD symptoms is likely to significantly reduce the number of cases identified (Bird, 1999) and decrease the risk that children with ADHD characteristics may be inappropriately labeled. Moreover, reliable measurement of specific areas of impairment associated with ADHD would allow areas of intervention to be targeted and meaningful improvements evaluated. Unfortunately, few impairment-related measures in child psychiatry have been developed, much less rigorously evaluated.

To date, the most commonly used measure purported to assess ADHD-related impairment is the Swanson, Kotkin, Agler, M-Flynn, and Pelham Scale (SKAMP; Swanson, 1992). The SKAMP specifically assesses context-bound behaviors critical to success in school settings, which is often the most problematic domain of functioning for children with ADHD. The 10 SKAMP items were initially developed as modifications of target behaviors used in specialized classroom management systems. As can be seen in the appendix, items are framed as “difficulties” and “problems” that would be expected to reflect ADHD-related classroom impairment, including the performance of academic tasks, following class rules, and interacting with peers and adults in the classroom. At the same time, several items appear quite similar to Diagnostic and Statistical Manual of Mental Disorders (4th ed; American Psychiatric Association, 1994) symptoms, raising questions about whether the SKAMP is indeed measuring anything other than school-based symptomatology.

The overlap of symptoms with impairment in ADHD measures may also reflect a conceptual overlap in these constructs. That is, the frequency of “often” defining *DSM-IV* symptoms is typically determined by clinicians based on parent and teacher reports of difficulties or impairments that arise because of ADHD behaviors. This is understandable given the lack of guidance provided by the *DSM-IV* for assessing impairment and underscores how impairment is embedded in the definition of symptoms. As both symptoms and impairment are required for *DSM-IV* diagnosis, samples of clinically diagnosed children are unlikely to demonstrate independence of these constructs. Theoretically, however, a child may manifest all the symptoms of ADHD but live in an environment where these create no difficulties and do not meet criteria for diagnosis. Thus, it would be helpful to clinicians to have tools to better distinguish symptoms from impairment.

The SKAMP has been used most often in small clinical samples to assess dosing and delivery strategies of stimulant medication (Greenhill et al., 2003; Swanson et al., 2002; Wigal, Gupta, Guinta, & Swanson, 1998). In this context, it has demonstrated sensitivity to treatment effects, with scores covarying across drug conditions and doses. Ratings conducted across two comparable drug conditions each day were moderately correlated, which suggests some reliability, and correlations with the Conners and IOWA I/O symptom rating scales were high ($r = .50 - .84$). These latter data were provided by Wigal et al. (1998) as evidence of concurrent validity of the SKAMP, although these correlations suggest that the SKAMP may be in large part redundant with a measure of symptoms. It was also used in the initial titration trial of the National Institute of Mental Health (NIMH) collaborative multisite Multimodal Treatment Study of Children with ADHD (MTA; Greenhill et al., 2001). Interestingly, SKAMP items were combined with a measure of the 18 *DSM-IV* ADHD symptoms, the Swanson, Nolan, and Pelham (SNAP) questionnaire (Swanson, 1992), due to high correlations and evidence from principal components analysis that they could be considered one factor. Thus, it remains unclear whether the SKAMP measures anything clinically distinct from ADHD symptoms.

The only published community sample evaluating the SKAMP was based on 109 predominantly Caucasian second-fifth graders from a school for children of military personnel (McBurnett, Swanson, Pfiffner, & Tamm, 1997). Principal components analysis in this study

identified two factors, labeled “Attention” and “Department,” which accounted for 71% of the variance in the items. Internal consistency for the total scale (.94) and the Attention (.95) and Department (.85) factors was high. However, the two subscales were moderately related to each other ($r = .53$), and strongly related to the SNAP ($r = .84$ for Inattention and .89 for Department with Hyperactivity/Impulsivity, respectively). The authors nonetheless report some support for what they identify as the divergent validity of the two factors (McBurnett et al., 1997). That is, the Attention scale was related to teacher ratings of Academic Competence on the Social Skills Rating Scale (Gresham & Elliott, 1990) and achievement scores on a standardized group test. The Department scale was related to teacher ratings of conduct problems on the Revised Behavior Problem Checklist (Quay & Peterson, 1983); the Conners, Loney, and Milich scale (CLAM; Loney & Milich, 1982); and the SNAP (Swanson, 1992) as well as negative peer nominations.

In sum, the SKAMP is commonly used for monitoring change in ADHD-related functioning in research settings and is supported by preliminary psychometric data in small clinical samples. However, its application in schools and community samples is severely limited by a lack of basic psychometric data, particularly from any samples large enough to provide norm-referenced data, which is necessary to differentiate normal variability in school functioning from that which may be associated with negative outcomes. Therefore, the present study addresses five specific aims with a large and diverse sample of public elementary school children: (a) examine the SKAMP's reliability; (b) evaluate its factor structure; (c) evaluate the convergent and discriminant validity of the SKAMP relative to the SNAP, an ADHD symptom measure; (d) examine the validity of the SKAMP in predicting concern and diagnostic criteria; and (e) provide normative data, including information on score variations by race, gender, age, and poverty.

Method

Subjects

Study procedures, including informed consent, were approved by the Institutional Review Board of the University of Florida and the school district research director. Participants were drawn from a longitudinal study designed to produce a representative sample of students at high risk for ADHD. Details of the study design are described elsewhere (Bussing, Zima, Gary, & Garvan, 2003). School registration records identified 12,009 students enrolled in kindergarten through fifth grade in a diverse north-central Florida public school district during the academic year 1998–1999. Of these, 3,251 students were selected for Phase 1 ADHD risk screening using a gender-stratified random sampling design, such that girls were oversampled by a margin of two to one to ensure adequate representation of girls with ADHD symptoms for subsequent study phases. Only one child per household was eligible for Phase 1 selection to ensure participant independence. Children were eligible for the study if they lived in a household with a telephone, were not receiving special education services for mental retardation or autism, and were from Caucasian or African American backgrounds. Children from other racial or ethnic backgrounds (e.g., Hispanic, Asian) were excluded because they composed less than 5% of the total student population in the school district.

Telephone contact was established with 63% of the selected sample ($n = 2,035$), and the respondents were primarily mothers. The remaining 37% were classified as unreachable due to nonworking phone numbers or because no contact could be made with multiple call attempts. Of those who could be reached, 79% ($n = 1,615$) agreed to participate, and 96% ($n = 1,549$) gave permission to obtain teacher ratings. Teachers returned 1,205 completed questionnaires (78% participation rate). Teacher questionnaire completion was slightly higher for economically advantaged children, (78% versus 72%), $\chi^2(1, 1613) = 7.82, p < .01$, Caucasians (76% versus 71%), $\chi^2(1, 1613) = 5.66, p < .05$, and children in the lower grades (77% versus

72%), $\chi^2(1, 1613) = 4.71, p < .05$, than for their disadvantaged, African American, or higher grade peers.

Phase 1—Interviews included inquiries into the child's health status, parental knowledge and attitudes about ADHD, a structured ADHD detection and service use assessment, and behavior ratings. As part of this structured interview, parents were asked whether there had been any general concerns that their child might have an emotional or behavioral problem; whether they or school staff suspected that their child had ADHD, attention deficit disorder, ADD, attention deficit, or hyperactivity; and whether their child had ever had a professional evaluation for ADHD. Of the 1,205 children with completed teacher ratings, 7% ($n = 89$) had reportedly received a professional ADHD diagnosis and were labeled “Diagnosed ADHD.” For 140 children (12%), either the parents or school staff had voiced a suspicion of ADHD, but no diagnostic assessment had been obtained; these children were labeled “Suspected ADHD.” For this study, the diagnosed and suspected children ($n = 229, 19\%$) together were classified as “ADHD-Specific Concern.” For another 332 children (28%), parents and/or school staff had voiced concerns about the child's emotions or behavior without suspicion or diagnosis of ADHD, and these children were classified as “Nonspecific Concern.” The remaining 644 children (53%) were captured in the category of “No Concern.”

Phase 2—Of the 1,205 Phase 1 children with teacher ratings, 266 were eligible for Phase 2 based on the presence of an ADHD-Specific Concern ($n = 229$) or a Nonspecific Concern with any parent-rated elevation above 2 standard deviations (SD) for age and gender ($n = 37$) on the Swanson, Nolan, and Pelham–IV (SNAP-IV) questionnaire (Swanson et al., 2001), an ADHD symptom measure. Of these, 190 (71%) participated, approximately 1 year after Phase 1. Phase 2 consisted of diagnostic interviews, self-report measures, and services assessments during home- or community-based personal interviews and collection of written permission for the release of school records.

Phase 3—All Phase 2 families were eligible to participate in Phase 3 (1 year after Phase 2) except those who had moved out of the school district. Of the eligible children, 156 had completed teacher ratings used for the present analyses, and 190 had information on discipline referrals. The parent interview included a structured ADHD detection and service use assessment, functional impairment ratings (Columbia Impairment Scale [CIS; Bird et al., 1993]), and behavior ratings.

Phase 4—All Phase 3 families were eligible to participate in Phase 4 (3 years after Phase 3) except families who had moved out of the district. Of the eligible families, 70% participated, yielding 106 Phase 4 participants for whom the Vanderbilt ADHD Diagnostic Parent Rating Scale (VADPRS) was collected. Vanderbilt ADHD Diagnostic Teacher Rating Scale (VADTRS) data were obtained for 73 of these adolescents.

Measures

SKAMP (Swanson, 1992)—The SKAMP is a 10-item scale designed to assess impairment associated with specific context-bound ADHD classroom behaviors. Teachers rate the severity of 10 items (6 for attention, such as “difficulty getting started on classroom assignments”; and 4 for deportment, such as “difficulty remaining quiet according to classroom rules”) on a 4-point scale: 0 = *not at all*, 1 = *just a little*, 2 = *pretty much*, to 3 = *very much*. It should be noted that subsequent versions of the SKAMP have been developed, including one with a 7-point scale and the addition of an individualized write-in item. This study examined the original version of the SKAMP (see appendix), which is sometimes embedded in the SNAP-IV. Teachers were asked to base their ratings on observations of the student over the previous 4 weeks.

SNAP-IV (Swanson et al., 2001)—The MTA version of the SNAP-IV was used to obtain symptom ratings from two sources, parents and teachers. The 26 items of the MTA SNAP-IV include the 18 ADHD symptoms (9 for inattentive, such as “often does not seem to listen when spoken to directly” and 9 for hyperactive/impulsive, such as “often fidgets with hands or feet, squirms in seat”) and 8 oppositional defiant disorder (ODD) symptoms, such as “often loses temper,” specified in the DSM-IV. Items are rated on a 4-point scale from 0 = *not at all* to 3 = *very much*. Typically, average rating-per-item (ARI) subscale scores for both parent and teacher scales are calculated for the inattention, hyperactivity/impulsivity, and opposition/defiance domains, resulting in six SNAP-IV ARI scores. Coefficient alphas for parent and teacher ratings calculated for the combined 26 items were .94 and .97, respectively, in this study.

Vanderbilt ADHD rating scale (VADPRS, VADTRS), parent and teacher versions (Wolraich, Feurer, Hannah, Baumgaertel, & Pinnock, 1998; Wolraich, Lambert, et al., 2003)—The Vanderbilt scale is a DSM-IV-based measure with parallel parent and teacher forms that include items measuring symptoms of ADHD, ODD/conduct problems, and anxiety/depression. For the present study, only the items assessing “performance” or impairment were examined. Parent and teacher forms each include 8 items on a 1–5 Likert scale assessing Academic Performance (e.g., reading, math, and written language) on a scale from “problematic” to “above average.” The VADTRS also includes items assessing classroom behavior and academic performance (relationships with peers, following directions/rules, disrupting class, assignment completion, and organizational skills), whereas the VADPRS includes items evaluating relationships with peers, siblings, and parents and participation in organized activities.

Diagnostic Interview Schedule for Children, Parent Version (NIMH DISC-IV-P; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000)—For Phase 2 participants, diagnoses of ADHD, ODD, and conduct disorder (CD) were made using the DISC-IV-P, which uses criteria contained in the DSM-IV and inquires about symptoms and impairment in both home and school settings. We used the standard DISC impairment algorithm, which requires moderate impairment in at least one area of functioning related to ADHD symptoms, as judged by the parent respondent. Impairment on the DISC is defined by the degree to which the symptoms have (a) caused distress to the child; (b) affected relations with caregivers, family, friends, or teachers; or (c) affected school functioning.

In its earlier versions, the DISC was shown to have moderate to substantial test–retest reliability and internal consistency (Fisher et al., 1993; Jensen et al., 1995; Piacentini et al., 1993). Cronbach's alpha for the DISC-P ADHD module is .93 in a referred sample (Wolraich et al., 2003). Despite its greater length and complexity, the test–retest reliability of the DISC-IV-P compares favorably with the earlier versions (Shaffer et al., 2000).

Columbia Impairment Scale (Bird et al., 1993)—Global impairment was assessed with the 13-item parent version of the CIS. Parents indicate how much of a problem they think the child has with, for example, getting along with his or her mother or with getting involved in activities. Items are scored on a Likert scale ranging from 0 = *no problem* to 4 = *a very big problem*. Two items tap into functioning relevant to the school setting (namely, behavior at school and schoolwork); however, factor analysis has suggested a single domain impairment score, and specific subscales have therefore not been identified (Bird et al., 1993). CIS authors also reported high internal consistency and test–retest reliability as well as significant correlations with clinician and parent ratings of child impairment. In this study, Cronbach's alpha was .86.

Sociodemographic characteristics and school services—Information about gender, age, race, grade level, special education services, and lunch subsidy status was obtained from school district administrative records. Table 1 presents this data for each phase of the study. Due to planned oversampling described below, there were almost twice as many girls as boys in the Phase 1 sample, although these numbers are generally equivalent for Phase 2 and 3. African American children composed a significant minority of the sample, which reflects school district demographics but is higher than expected for the county. Child lunch status, identified as subsidized or nonsubsidized based on federal government guidelines involving family income, was used as an indicator of socioeconomic status (SES), with subsidized lunch corresponding to lower SES. The Hollingshead (1975) Four Factor Index, which ranges from 8 (lowest social strata) to 66 (highest strata), based on parental education and occupation, was also calculated.

Discipline referrals—As with sociodemographic data, we also collected information about discipline referrals from computerized school district records, which were reported to the state for educational accountability purposes. Associated with higher parent and teacher ratings of disruptive behavior (Rusby, Taylor, & Foster, 2007), discipline referrals have been examined previously as an indicator of children's behavioral school adjustment (Kim, Kamphaus, & Baker, 2006). In this study, we calculated the cumulative number of disciplinary referrals a student received between the Phase 1 and the Phase 3 interviews as an indicator of behavioral impairment in the school setting. The mean number of referrals for the 156 students on which these data were available was 4.26 ($SD = 10.77$), with a range of 0 to 95 over the 2-year period examined (see Table 1). However, over half of the sample had no referrals, reflecting a highly skewed variable, as might be expected.

Data Analysis

Distribution characteristics of the SKAMP were nonnormal, as is often found for Likert scale data in nonclinical samples. Because the use of parametric methods such as *t* tests and Pearson correlations depends on assumptions of normally distributed data for validity, we used distribution-free methods, such as Wilcoxon rank sum tests and Spearman correlations, which are efficient and robust, except when comparison to parametric values was considered helpful. Most of our analyses were based on the entire sample of 1,205 children, but we have indicated when this varied due to the phase of data collection. It should also be noted that we examined the SKAMP in relation to several parent-completed measures, expecting that relations may be lower due to rater effects but believing that full exploration of the SKAMP's validity or lack thereof would be helpful.

Our analyses were adjusted for sample design, including the oversampling of girls and for differential response using analytic weights computed in a procedure outlined by Aday (1996). This procedure effectively weighted the sample to be more representative of the target population. For example, because girls were sampled in a two-to-one ratio, the analytic weight for a girl is less than the analytic weight for a boy. This weighting process allowed us to nearly match the representation of various subgroups in our sample (e.g., Caucasian girls receiving subsidized lunches) with the subgroup percentages for the entire population. Without these statistical adjustments, the overrepresentation of girls in the sample would skew the mean total sample data.

In the first stage of weight development, an expansion weight (the inverse of the selection probability) was computed for each subject that depended on child gender and the number of eligible children in a household. In the second stage of weight development, 12 weighting classes were formed based on factors where significant differential response was noted, which included race, lunch subsidy status, and special education service status (Cox & Cohen,

1985). To adjust for differential response rates, the expansion weight was divided by the response rate within each weighting class to form a response-adjusted weight. In the third stage of weight development, a relative weight was constructed by dividing each response-adjusted weight by the mean response-adjusted weight. This scaling step effectively downweighted the number of subjects to equal the actual sample size. The final weight was obtained after trimming the extreme (lower 1% and upper 99%) values of the relative weights and uniformly redistributing the values so that the actual sample size was preserved.

In accordance with our five aims evaluating the psychometric properties of the SKAMP, we first examined reliability for the entire sample through item analysis and internal consistency (Aim 1). Item–total correlations were calculated using Spearman correlation coefficient r_s due to the noncontinuous and nonnormal nature of the data. To determine scale reliability, we fit a confirmatory one-factor congeneric matrix measurement model to a scaled covariance matrix of polychoric correlations among the SKAMP items, a method recommended by Rowe and Rowe (1997) that is more rigorous than coefficient alpha.

Our second aim, examining the SKAMP's factor structure, was accomplished by using exploratory factor analysis (EFA) methods, followed by confirmatory factor analyses (CFA). EFA was performed using a split-sample technique with polychoric correlation matrices. To avoid the common problem of overfactoring due to use of liberal statistical criteria for EFA, we followed recommendations by Frazier and Youngstrom (2007). They identified two procedures, Horn's (1965) parallel analysis (HPA) and the minimum average partial (MAP) analysis, as gold standard methods for identifying the true number of existing factors.

Although infrequently used, HPA and MAP have demonstrated better accuracy than more commonly used factor analysis procedures (Velicer, Eaton, & Fava, 2000; Zwick & Velicer, 1986). HPA compares the eigenvalues obtained from principal components analysis of the observed correlation matrix to eigenvalues obtained from a randomly generated correlation matrix. Components from the observed data that have eigenvalues larger than the upper bound of the 95% confidence interval of randomly generated eigenvalues are retained. MAP (Velicer, 1976) examines successive partial correlation matrices in which the average squared correlation of the observed correlation matrix is computed, and successive components resulting from principal components analysis are partialled from the original matrix until the minimum average squared partial correlation is obtained, thereby indicating the number of components to retain.

The CFA was conducted using the CALIS procedure in SAS. Polychoric correlation matrices were used to fit the confirmatory factor models given the ordinal nature of the SKAMP data. A Box–Cox power transformation was then used to yield data with approximately normal score distributions (Box & Cox, 1964). To assist in interpreting results of the above methods, we performed the Schmid-Leiman procedure to examine the proportion of general variance and specific variance accounted for by the factors. We also examined incremental validity of the two SKAMP subscales through sequential regression analyses of discipline referrals and Vanderbilt ratings, adding each subscale to the total score.

Our third aim was to examine the convergent and discriminant validity of the SKAMP in relation to the SNAP. Given questions about whether the SKAMP may be conceptually and/or statistically different from ADHD symptom measures, we examined correlations between the SKAMP and SNAP and measures of impairment and also conducted multiple and Poisson analyses to determine how much the SKAMP adds to the prediction of outcomes above and beyond that provided by the SNAP alone. Our outcomes included (a) disciplinary referrals between the Phase 1 and Phase 3 interview ($n = 156$) modeled in the form of counts, (b) parental reports of child functioning obtained on the CIS at the Phase 3 interview ($n = 156$), (c) teacher

ratings of school functioning at Phase 4 on the VADTRS ($n = 73$), and (d) parent ratings of academic and social functioning on the VADPRS at Phase 4 ($n = 106$).

Correlations were examined using a nonparametric statistic, Kendall's τ correlation coefficient (Newson, 2002), due to the ordinal nature of SKAMP data. However, because the Spearman rank-order correlation coefficient r_s is more commonly used to produce estimates resembling the amount of explained variability, we report both estimates. Kendall's τ is comparable to Spearman's r_s with regard to the underlying assumptions and statistical power; however, Kendall's τ has substantial advantages over the Spearman coefficient because it has better statistical properties and allows testing of whether two correlations are significantly different as indicated by z scores. Formulas to convert estimates for the Kendall's τ into other correlational indices have been developed (Walker, 2003), showing that Kendall's τ is usually smaller than Spearman's r_s .

As our fourth aim, we assessed the validity of the SKAMP in predicting concurrent parent concern level at Phase 1 and the presence or absence of ADHD diagnoses on the DISC-IV (at Phase 2) using the GLM procedure and adjusting for multiple comparisons with the Tukey-Kramer method. Our fifth and final aim, providing normative data for the SKAMP, was addressed by first exploring whether SKAMP scores differed across subgroups. To do this, we conducted a multivariate analysis of variance (MANOVA) using the two subscale scores as the dependent variables and gender, race, poverty, and age as predictors. We calculated Cohen's d values, defined as the standardized mean difference between groups (Cohen, 1988), using the Wilcoxon rank-sum test to explore the meaningfulness of differences among the subgroups (Field & Hole, 2003).

Results

Aim 1: Reliability

Spearman correlations for the Department ratings ranged from .67 ("problems in interaction with staff") to .70 ("difficulty staying seated according to classroom rules") and for the Attention ratings, from .71 ("problems in accuracy or neatness of written work in the classroom") to .88 ("difficulty staying on task for a classroom period"). Internal consistency estimates were high, with reliabilities of .98 for overall SKAMP scores, .96 for Department, and .95 for Attention.

Aim 2: Factor Structure

Results of the exploratory factor analyses suggested three factors in the structure, with Items 1, 2, 7, 8, 9, and 10 comprising the first factor (consistent with McBurnett's Attention subscale), Items 5 and 6 from the Department scale comprising the second factor, and Items 3 and 4 comprising a third factor reflecting interpersonal impairment at school. However, this latter factor was minor and explained only 3% of variation. Subsequent HPA results also suggested two main and one minor factor, whereas MAP analysis suggested two factors, with Items 1, 2, 7, 8, 9, and 10 loading on the Attention factor identified by McBurnett et al. (1997) and Items 3, 4, 5, and 6 on the Department factor. Based on the Schmid-Leiman procedure, we found that 25% of variance was accounted for in the two-factor model, and 18% of variance was accounted for in the three-factor model. Thus, although conceptually interesting, the third factor did not appear to provide meaningful additional variance.

To further evaluate a two-factor model, we conducted incremental validity analyses based on discipline referrals, which indicated that addition of the two SKAMP subscales significantly improved the prediction of referrals beyond the total SKAMP score [*log likelihood ratio test statistic* = 8.25, has χ^2 degrees of freedom (df) = 1, $p = .004$]. However, incremental validity

analyses based on the Vanderbilt scale did not indicate a significant increase in the amount of variance explained by the two SKAMP subscales above that explained by the SKAMP total score, $R^2 = .10$ for VADPRS with total SKAMP alone and $R^2 = .11$ with the addition of either SKAMP subscale and $R^2 = .23$ for VADTRS with total SKAMP alone and $R^2 = .23$ with addition of either SKAMP subscale. Thus, partial support was found for the two subscales in these analyses.

Overall, EFA and incremental validity analyses support the SKAMP two-factor model as described by McBurnett et al. (1997). However, given the possibility of a third factor, we also included a three-factor model in the subsequent CFA, in addition to a one-factor model and a two-factor model. We examined multiple goodness-of-fit indices, recognizing there is no single, generally accepted model of fit index and that large sample sizes can affect indices. As shown in Table 2, indices for the two-factor solution based on the overall sample ($n = 1,205$) were consistently better than for the one-factor solution but did not fall in the acceptable range for all indices. More specifically, the root mean square error of approximation (RMSEA), considered a better indicator of model fit when there is a substantial relation among factors (Rigdon, 1996), as in the present case, was above .05 for all models. Nonetheless, the two-factor model had better fit indices than the one-factor model (difference of $\chi^2 = 926.72$, $p < .0001$; AIC smaller). Even though the three-factor model had better fit than the two-factor model on these indices as well, this finding was outweighed by the results obtained in the EFA, the HPA, the MAP, and the sequential analysis, showing stronger support for the two-factor model.

The two SKAMP factors were highly related (loading between two domains = .85), as is often found between the Inattention and Hyperactivity/Impulsivity domains on ADHD symptom ratings. Nonetheless, given theoretical and conceptual views of ADHD as having two distinct domains, we thought it important to examine the validity of the SKAMP for both factors as well as the overall score. Therefore, all additional psychometric evaluations were conducted in this manner.

Aim 3: Convergent and Discriminant Validity in Relation to the SNAP

Table 3 shows correlation indices for the overall SKAMP and its Attention and Department subscales with parent and teacher SNAP-IV ratings and subscales (Inattention, Hyperactivity/Impulsivity, ODD) for the full sample ($n = 1,205$). As can be seen, all correlations for the two measures across subscales are statistically significant, with significantly stronger correlations for the SKAMP with the teacher SNAP than the parent SNAP ($z = 21.88$, $p < .0001$ for the total scores), likely reflecting rater effects. The SKAMP Attention subscale correlated more strongly with the SNAP-IV teacher Inattention ratings than with teacher Hyperactivity/Impulsivity ($z = 17.84$, $p < .0001$) and with ODD ratings ($z = 20.99$, $p < .0001$), whereas the Department subscale correlated more strongly with the latter two ($z = 4.56$, $p < .0001$ for Hyperactivity/Impulsivity; $z = 2.70$, $p < .01$ for ODD) than with Inattention ratings.

A similar but not entirely consistent pattern of results was found for correlations of the SKAMP subscales with parent SNAP ratings. More specifically, the Attention subscale correlated more strongly with parent SNAP-IV Inattention than with parent Hyperactivity/Impulsivity ($z = 4.12$, $p < .0001$) and ODD ($z = 7.66$, $p < .0001$) ratings. However, the Department subscale correlated more strongly with both parent SNAP-IV Inattention and Hyperactivity/Impulsivity than parent ODD ratings ($z = 3.50$, $p < .0001$ for Inattention and $z = 4.19$, $p < .0001$ for Hyperactivity/Impulsivity). No differences were found between Department subscale correlations for parent SNAP-IV Hyperactivity/Impulsivity and Inattention.

As shown in Table 3, the total SKAMP score as well as the two subscales significantly predicted the number of future discipline referrals ($r_s = .41 - .46$, $p < .001$), and adding the SKAMP to the SNAP subscales significantly improved the SNAP's prediction [*log likelihood ratio test*

statistic = 448.64, has $\chi^2 df = 2, p < .001$]. The SKAMP also predicted teacher ratings of classroom impairment on the VADTRS ($r_s = .36 - .52, p < .001$) and parent ratings of impairment on the VADPRS ($r_s = .21 - .39, p < .01$) several years later. The amount of variance explained on the VADPRS increased from $R^2 = .17$ based on the SNAP subscales alone to $R^2 = .25$ when the SKAMP subscales were added, with $F(1, 101) = 5.61, p = .020$, for SKAMP Attention and $F(1, 101) = .01, p = .923$, for SKAMP Deportment. The corresponding findings for teacher-rated impairment on the VADTRS were $R^2 = .19$ versus $R^2 = .31$, with $F(1, 68) = 4.35, p = .041$, for SKAMP Attention and $F(1, 68) = 3.20, p = .078$, for SKAMP Deportment.

In contrast, SKAMP scores did not correlate with overall parent-reported CIS scores or increase the variance predicted by the SNAP alone. Because the CIS is very broad in nature and contains several items not reflective of a child's school functioning, we also assessed correlations between SKAMP scores and the two CIS items with school relevance (i.e., School Behavior, Schoolwork). Significant correlations were found between teacher SKAMP scores (summary and subscales) and School Behavior ($r_s = .31 - .35, p < .001$). Schoolwork problems were also correlated with teacher Attention scores ($r = .23, p < .01$), but not Deportment ratings ($r = .07, p > .1$). Adding SKAMP ratings to the SNAP significantly increased the amount of variance accounted for in the two CIS school items from $R^2 = .21$ to $R^2 = .24$, with $F(2, 151) = 3.10, p = .048$.

Aim 4: Validity in Predicting Concern and Diagnostic Criteria

As can be seen in Table 4, Attention scores for children in the ADHD-Specific Concern group (mean [M] = 1.51, $SD = 1.06$) were significantly higher than those for children in the Nonspecific Concern group ($M = 1.08, SD = .94, t = 6.31, p < .001$), and Attention scores for the latter were significantly higher than those for children in the No Concern group ($M = .46, SD = .65, t = 11.14, p < .001$). Similarly, Deportment scores for children in the ADHD Specific Concern group ($M = 1.11, SD = .97$) were significantly higher than those for children in the Nonspecific Concern group ($M = .73, SD = .76, t = 6.67, p < .001$), and Deportment scores for the latter were significantly higher than those for children in the No Concern group ($M = .32, SD = .54, t = 8.67, p < .001$). In contrast, for the Phase 2 group, SKAMP scores were unsuccessful in discriminating those children who met full criteria for ADHD on the DISC-IV by parent report from those who did not, with no differences between groups on either of the SKAMP subscales.

Aim 5: Normative Data

MANOVA indicated statistically significant differences by gender (Wilks's $\Lambda = .93; F(2, 1199) = 43.51, p < .0001$), race (Wilks's $\Lambda = .94; F(2, 1199) = 27.07, p < .0001$), and poverty, as defined by free and reduced lunch status (Wilks's $\Lambda = .97; F(2, 1199) = 16.95, p < .0001$), but not for age (Wilks's $\Lambda = .99; F(2, 1199) = 1.79, p = .167$). Cohen's d values for gender differences in SKAMP scores were small to moderate for overall (.52), Deportment (.45), and Attention (.50) ratings, with consistently higher ratings for boys than for girls. Cohen's d estimates for SKAMP score differences by race and by poverty were moderate for overall (.66, .61), Deportment (.60, .54), and Attention ratings (.63, .58), consistently showing higher teacher ratings for African American children and children living in poverty than Caucasian children and non-socioeconomically disadvantaged children. Therefore, we chose to present normative data by gender and race as well as the full sample (see Table 4).

Race was highly correlated with poverty in our sample (tetrachoric correlation $n = .82$), with 89% of African American children receiving free or reduced lunch (odds ratio = 21.07, 95% confidence interval [CI] = 14.54, 30.55). To establish whether race and social disadvantage independently predict SKAMP scores, we performed a multiple regression analysis using the GLM procedure. Race, poverty, and SES scores based on the Hollingshead (1975) index

emerged as independent predictors for SKAMP subscale scores, with beta estimates of .29 ($p < .0001$), .16 ($p < .01$), and $-.01$ ($p < .001$), respectively, for Dependent, and of .39 ($p < .0001$), .18 ($p < .01$), and $-.01$, ($p < .0001$) for Attention.

For exploratory purposes, we have also provided descriptive data stratified by race and poverty subgroups in Table 5. SKAMP total scores for socioeconomically disadvantaged Caucasian children ($M = .73$) are higher than for their non-socioeconomically disadvantaged Caucasian peers ($M = .43$, $t = 4.57$, $p < .001$) but lower than for African American socioeconomically disadvantaged children ($M = 1.16$, $t = -6.37$, $p < .001$). Of note, SKAMP Total scores for African American non-socioeconomically disadvantaged children ($M = .66$) did not differ significantly from those of non-socioeconomically disadvantaged Caucasians ($t = 1.38$, $p = .31$), although this may be due to the small sample size of the former ($n = 37$). A similar pattern was noted for each of the SKAMP subscales.

Discussion

The goal of this study was to evaluate the SKAMP's psychometric properties using multiple statistical approaches in a large, epidemiologically derived community sample. Given the lack of basic psychometric data, we examined five specific aims pertaining to (a) the SKAMP's reliability; (b) factor structure; (c) validity in relation to the SNAP, a symptom measure; (d) validity in predicting diagnostic criteria; and (e) we provide normative data, including information on race, gender, age, and poverty. Overall, this appears to be a reliable measure with two factors, which relates highly to the SNAP but also provides some unique variance in explaining future functional outcomes.

Results suggest the presence of at least two factors and possibly a third, which is small (e.g., two items reflecting interpersonal interactions). Although we considered presenting a three-factor solution, the latter had insufficient reliable variance to provide meaningful improvements in clinical decision making as currently constructed. Overall, CFA data support McBurnett et al.'s (1997) proposed two-factor model, although fit indices were most acceptable for boys. We suggest maintaining a distinction between these two related but separate constructs in the assessment and treatment of ADHD-related impairment. By examining the SKAMP Attention and Dependent factors separately, a more complete picture of a child's impaired functioning may emerge to facilitate more appropriate and effective treatment. However, future work on the application of the two-factor model for different sociodemographic subgroups would be informative. Future studies might also consider additional items that could contribute to more reliable assessment of interpersonal interactions as an area of classroom impairment.

As expected, the SKAMP was found to be related to both parent and teacher versions of the SNAP-IV ($r = .93$ and $.79$ for Inattention and Hyperactivity/Impulsivity). This convergence was expected given the overlap in domain items, particularly between academic functioning and symptoms of inattention, although we will note that SKAMP items ask about functioning related to symptoms captured in the SNAP. To some extent, this highlights the interdependence of symptoms and impairment, as currently defined in the *DSM-IV*.

Given the theoretical and statistical overlap between the SKAMP and SNAP, we felt it was important to examine the SKAMP's convergent and discriminant validity relative to the SNAP. Our results indicate that SKAMP scores add a statistically significant albeit modest amount to the prediction of future school discipline referrals as well as parent and teacher ratings of impairment several years later. However, the meaningfulness of this contribution with regard to the SKAMP's practical utility remains unclear. The extent to which the SKAMP predicts supplemental variance relative to other measures of behavior and functioning is also unknown.

Future research should examine the predictive and incremental validity of the SKAMP using other school-related impairment criteria, such as achievement scores and referrals for special education services to address its validity as an impairment measure.

We found relatively strong correlations between the SKAMP and school disciplinary referrals 2 or more years later as well as parent and teacher ratings of academic performance, school functioning, and social relations, as measured by the Vanderbilt scale 5 or more years later. The lack of relationship between the SKAMP and the overall CIS may be related to characteristics of that particular measure, which is more general than the other outcomes examined. However, these data provide evidence of predictive validity for future school-related outcomes.

Although the SKAMP may reflect setting-specific dysfunction, it is not necessarily etiology specific. That is, it may pick up on similar behavioral manifestations due to other causes besides ADHD that affect children's classroom competencies, such as anxiety, mood, and other disruptive disorders. Indeed, we found that ratings of oppositional behavior on the SNAP correlate with SKAMP scores, although relations were not as strong as they were between the SKAMP and ratings of Inattentive and Hyperactive/Impulsive behavior subscale. On the other hand, some support for ADHD specificity was found, in that SKAMP scores were higher for those children with ADHD-specific concerns (previously diagnosed or suspected by parents or teachers) than for those with nonspecific concerns.

However, the SKAMP was not able to predict later diagnosis of ADHD on the DISC-IV. The most parsimonious explanation for this is informant variance, which is reported to be equal or stronger to trait (symptom) effects on other ADHD rating scales (Gomez, Burns, Walsh, & de Moura, 2003). That is, the SKAMP was completed by teachers while the DISC was completed by parents, who observe children in different settings and are likely to have different views of their functioning. The finding that the SKAMP does discriminate children based on ADHD-specific concerns may also reflect that concern level was based on either parent or teacher concern, unlike the DISC, which was based on parent report only.

Although the SKAMP does not purport to generate categorical diagnostic decisions, we thought it was important to examine whether any differences might exist on the SKAMP by age, gender, and ethnicity. We found gender and ethnic differences consistent with previous research on ADHD symptom measures completed by teachers (DuPaul et al., 1997; Epstein, March, Conners, & Jackson, 1998; Reid et al., 1998). That is, boys were rated as having more classroom ADHD-related impairment than girls, with small to moderate effect sizes. In addition, even when controlling for SES and poverty level, African American students were rated as more impaired than Caucasian students, with moderate effect sizes ($d = .56 - .61$). Our presentation of normative data for the total sample and separately by gender and ethnicity allows the potential user to calculate severity of SKAMP impairment relative to the general population, while also considering impairment in a race-specific context, as recommended by Collett et al. (2003). The lack of any age effects may be related to there being few SKAMP items that assess overt hyperactive behaviors, which have been shown to decrease over time (Hart, Lahey, Loeber, Applegate, & Frick, 1995). Age effects were also absent in this sample on the SNAP, a measure of ADHD symptoms (Bussing et al., in press). This may reflect some unique characteristic of the present sample, although other ADHD studies have also reported small to negligible age effects (Conners, 1997).

We acknowledge limitations of our data for interpreting ethnic differences in this study. We did not obtain data on teacher ethnicity or on contextual factors related to schools or classrooms, which could account for possible sources of bias (Epstein et al., 2005; Reid, Casat, Norton, Anastopoulos, & Temple, 2001). In addition, the teacher response rate was lower for African

American and lower SES students. Elevated ratings for African Americans on rating scales such as the SKAMP may also be due to differences in score distributions or instrument bias, as suggested by Reid et al. (1998). We considered presenting results by poverty status instead of ethnicity; however, stratification of scores by poverty status is less likely useful clinically, as poverty itself may vary over time, complicating its use as an assessment factor. Nonetheless, we recognize that elevated scores for African American children on the SKAMP may be due to the association of ethnic status to poverty rates, a risk factor for ADHD (Biederman et al., 1995). Thus, caution in interpreting results with African American children may be indicated. Further research investigating the meaning of race differences and poverty effects on the SKAMP is indicated, particularly given exploratory data suggesting that African American non-socioeconomically disadvantaged students may not differ from Caucasian non-socioeconomically disadvantaged students on the SKAMP.

The data for this study were obtained from one school district, which although diverse and drawn from a county that was sociodemographically representative of the state at the time, includes a higher percentage of socioeconomically disadvantaged and African American children (30%) than is nationally representative (U.S. Census Bureau, 2000). Moreover, we were unable to examine children of other minorities or those from multiracial backgrounds, limiting application of our results to those populations. However, our results are consistent with previous research examining the SKAMP in a predominantly Caucasian sample from a different region of the United States (McBurnett et al., 1997), which increases confidence in the generalizability of these findings.

In sum, the SKAMP has many advantages, as outlined by Collett et al. (2003), including short administration time, minimal respondent burden, feasibility as a Web-based instrument (Bhatara, Vogt, Patrick, & Doniparthi, 2006), and established sensitivity to treatment. This study provides comprehensive psychometric data from a large, sociodemographically diverse sample and supports its reliability and validity, including retention of two factors. Examination of the SKAMP in relation to the SNAP indicates initial support for some meaningful differences, although determining whether the SKAMP can indeed be considered a measure of ADHD-related impairment will require further work examining the SKAMP in relation to a number of well-established measures of symptoms and impairment.

Appendix

Appendix SKAMP

Instructions: Please indicate the answer that best describes this child in the classroom during the last 4 weeks. Select only one response for each question.

	Not at All	Just a Little	Pretty Much	Very Much
1. Difficulty getting started on classroom assignments	0	1	2	3
2. Difficulty staying on task for a classroom period	0	1	2	3
3. Problems in interactions with peers in the classroom	0	1	2	3
4. Problems in interactions with staff (teacher or aide)	0	1	2	3
5. Difficulty remaining quiet according to classroom rules	0	1	2	3
6. Difficulty staying seated according to classroom rules	0	1	2	3

	Not at All	Just a Little	Pretty Much	Very Much
7. Problems in completion or work on classroom assignments	0	1	2	3
8. Problems in accuracy or neatness or written work in the classroom	0	1	2	3
9. Difficulty attending to an activity or discussion of the class	0	1	2	3
10. Difficulty stopping and making transition to the next period	0	1	2	3

Source: Swanson (1992).

References

- Aday, LA. Designing and conducting health surveys. Jossey-Bass; San Francisco: 1996.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. Vol. 4th ed.. Author; Washington, DC: 1994.
- Bhatara V, Vogt H, Patrick S, Doniparthi E. Acceptability of a Web-based attention-deficit/hyperactivity disorder scale (T-SKAMP) by teachers: A pilot study. *Journal of the American Board of Family Medicine* 2006;19(2):195–200. [PubMed: 16513909]
- Biederman J, Milberger S, Faraone SV, Kiely K, Guite J, Mick E, et al. Family–environment risk factors for attention-deficit hyperactivity disorder. A test of Rutter's indicators of adversity. *Archives of General Psychiatry* 1995;52(6):464–470. [PubMed: 7771916]
- Bird, HR. The assessment of functional impairment.. In: Shaffer, D.; Lucas, CP.; Richters, JE., editors. *Diagnostic assessment in child and adolescent psychopathology*. Guilford; New York: 1999. p. 209-229.
- Bird HR, Shaffer D, Fisher P, Gould MS, Staghezza B, Chen JY, et al. The Columbia Impairment Scale (CIS): Pilot findings on a measure of global impairment for children and adolescents. *International Journal of Methods in Psychiatric Research* 1993;3:167–176.
- Box GEP, Cox DR. An analysis of transformations. *Journal of the Royal Statistical Society* 1964;26:211–252.
- Bussing R, Fernandez M, Harwood M, Hou W, Garvan CW, Eyberg SM, et al. Parent and teacher SNAP-IV ratings of attention-deficit/hyperactivity disorder: Psychometric properties and normative ratings from a school district sample. *Assessment*. (in press)
- Bussing R, Zima BT, Gary FA, Garvan CW. Barriers to detection, help-seeking, and service use for children with ADHD symptoms. *Journal of Behavioral Health Services and Research* 2003;30(2):176–189. [PubMed: 12710371]
- Cohen, J. Statistical power analysis for the behavioral sciences. Vol. 2nd ed.. Academic Press; New York: 1988.
- Collett BR, Ohan JL, Myers KM. Ten-year review of rating scales. V: Scales assessing attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* 2003;42(9):1015–1037. [PubMed: 12960702]
- Conners, CK. *Conners rating scales: Revised technical manual*. Multi-Health Systems; North Tonawanda, NY: 1997.
- Cox, BG.; Cohen, SB. *Methodological issues for health care surveys*. Marcel Dekker; New York: 1985.
- DuPaul GJ, Power TJ, Anastopoulos AD, Reid R, McGoey KE, Ikeda MJ. Teacher ratings of attention-deficit/hyperactivity disorder symptoms: Factor structure and normative data. *Psychological Assessment* 1997;9(4):436–444.
- Epstein JN, March JS, Conners CK, Jackson DL. Racial differences on the Conners Teacher Rating Scale. *Journal of Abnormal Child Psychology* 1998;26(2):109–118. [PubMed: 9634133]

- Epstein JN, Willoughby M, Valencia EY, Tonev ST, Abikoff HB, Arnold LE, et al. The role of children's ethnicity in the relationship between teacher ratings of attention-deficit/hyperactivity disorder and observed classroom behavior. *Journal of Consulting & Clinical Psychology* 2005;73(3):424–434. [PubMed: 15982140]
- Field, AP.; Hole, GJ. How to design and report experiments. Sage; Thousand Oaks, CA: 2003.
- Fisher PW, Shaffer D, Piacentini JC, Lapkin J, Kafantaris V, Leonard H, et al. Sensitivity of the diagnostic interview schedule for children, 2nd edition (DISC-2.1) for specific diagnoses of children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry* 1993;32(3):666–673. [PubMed: 8496131]
- Frazier TW, Youngstrom EA. Historical increase in the number of factors measured by commercial tests of cognitive ability: Are we overfactoring? *Intelligence* 2007;35(2):169–182.
- Gomez R, Burns GL, Walsh JA, de Moura MA. A multitrait–multisource confirmatory factor analytic approach to the construct validity of ADHD rating scales. *Psychological Assessment* 2003;15(1):3–16. [PubMed: 12674720]
- Greenhill LL, Swanson JM, Steinhoff K, Fried J, Posner K, Lerner M, et al. A pharmacokinetic/pharmacodynamic study comparing a single morning dose of Adderall to twice-daily dosing in children with ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry* 2003;42(10):1234–1241. [PubMed: 14560174]
- Greenhill LL, Swanson JM, Vitiello B, Davies M, Clevenger W, Wu M, et al. Impairment and deportment responses to different methylphenidate doses in children with ADHD: The MTA titration trial. *Journal of the American Academy of Child and Adolescent Psychiatry* 2001;40(2):180–187. [PubMed: 11211366]
- Gresham, FM.; Elliott, SN. Social skills rating systems manual. American Guidance Service; Circle Pines, MN: 1990.
- Hart EL, Lahey BB, Loeber R, Applegate B, Frick PJ. Developmental changes in attention-deficit/hyperactivity disorder in boys: A four-year longitudinal study. *Journal of Abnormal Child Psychology* 1995;23:729–750. [PubMed: 8609310]
- Hollingshead, AB. Four factor index of social class. Yale University, Department of Sociology; 1975. Unpublished manuscript
- Horn JL. A rationale and test for the number of factors in factor analysis. *Psychometrika* 1965;30:179–185. [PubMed: 14306381]
- Jensen P, Roper M, Fisher P, Piacentini J, Canino G, Richters J, et al. Test–retest reliability of the Diagnostic Interview Schedule for Children (DISC 2.1). Parent, child, and combined algorithms. *Archives of General Psychiatry* 1995;52(1):61–71. [PubMed: 7811163]
- Kim S, Kamphaus RW, Baker JA. Short-term predictive validity of cluster analytic and dimensional classification of child behavioral adjustment in school. *Journal of School Psychology* 2006;44:287–305.
- Loney, J.; Milich, R. Hyperactivity, inattention, and aggression in clinical practice.. In: Wolraich, M.; Routh, D., editors. *Advances in developmental and behavioral pediatrics*. Vol. 2. JAI; Greenwich, CT: 1982. p. 113-147.
- McBurnett K, Swanson JM, Pfiffner LJ, Tamm L. A measure of ADHD-related classroom impairment based on targets for behavioral intervention. *Journal of Attention Disorders* 1997;2(2):69–76.
- Newson R. Parameters behind “nonparametric” statistics: Kendall's tau, Somers' D, and median differences. *The Stata Journal* 2002;2(1):45–64.
- Piacentini J, Shaffer D, Fisher P, Schwab-Stone M, Davies M, Gioia P. The Diagnostic Interview Schedule for Children—Revised Version (DISC-R): III. Concurrent criterion validity. *Journal of the American Academy of Child and Adolescent Psychiatry* 1993;32(3):658–665. [PubMed: 8496130]
- Quay, HC.; Peterson, DR. Interim manual for the Revised Behavior Problems Checklist. Authors; Miami, FL: 1983.
- Reid R, Casat CD, Norton HJ, Anastopoulos AD, Temple EP. Using behavior rating scales for ADHD across ethnic groups: The IOWA Conners. *Journal of Emotional and Behavioral Disorders* 2001;9(4):210–218.

- Reid R, DuPaul GJ, Power TJ, Anastopoulos AD, Rogers-Adkinson D, Noll MB, et al. Assessing culturally different students for attention deficit hyperactivity disorder using behavior rating scales. *Journal of Abnormal Child Psychology* 1998;26(3):187–198. [PubMed: 9650625]
- Rigdon, EE. Proceedings of the Summer Educators' conference. American Marketing Association; Chicago: 1996. CFI vs. RMSEA: A comparison of two factor indices for structural equation modeling.; p. 37-38.
- Rowe KS, Rowe KJ. Norms for parental ratings on Conners' Abbreviated Parent-Teacher Questionnaire: Implications for the design of behavioral rating inventories and analyses of data derived from them. *Journal of Abnormal Child Psychology* 1997;25(6):425–451. [PubMed: 9468105]
- Rusby JC, Taylor TK, Foster EM. A descriptive study of school discipline referrals in first grade. *Psychology in the Schools* 2007;44:333–350. [PubMed: 17372623]
- Shaffer D, Fisher P, Lucas CP, Dulcan MK, Schwab-Stone ME. 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* 2000;39(1):28–38. [PubMed: 10638065]
- Swanson, JM. School-based assessments and interventions for ADD students. K.C. Publishing; Irvine, CA: 1992.
- Swanson JM, Gupta S, Williams L, Agler D, Lerner M, Wigal S. Efficacy of a new pattern of delivery of methylphenidate for the treatment of ADHD: Effects on activity level in the classroom and on the playground. *Journal of the American Academy of Child and Adolescent Psychiatry* 2002;41(11):1306–1314. [PubMed: 12410072]
- Swanson JM, Kraemer HC, Hinshaw SP, Arnold LE, Conners CK, Abikoff HB, et al. Clinical relevance of the primary findings of the MTA: Success rates based on severity of ADHD and ODD symptoms at the end of treatment. *Journal of the American Academy of Child and Adolescent Psychiatry* 2001;40(2):168–179. [PubMed: 11211365]
- U.S. Census Bureau. 2000 [March 23, 2007]. from <http://www.census.gov/main/www/cen2000.html>
- Velicer WF. Determining the number of components from the matrix of partial correlations. *Psychometrika* 1976;41(3):321–327.
- Velicer, WF.; Eaton, CA.; Fava, JL. Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components.. In: Goffin, RD.; Helmes, E., editors. Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy. Kluwer Academic; Norwell, MA: 2000. p. 41-71.
- Walker D. Converting Kendall's tau for correlational or meta-analytic analyses (SPSS). *Journal of Modern Applied Statistical Methods* 2003;2(2):525–530.
- Wigal SB, Gupta S, Guinta D, Swanson JM. Reliability and validity of the SKAMP rating scale in a laboratory school setting. *Psychopharmacology Bulletin* 1998;34:47–53. [PubMed: 9564198]
- Wolraich ML, Feurer ID, Hannah JN, Baumgaertel A, Pinnock TY. Obtaining systematic teacher reports of disruptive behavior disorders utilizing *DSM-IV*. *Journal of Abnormal Child Psychology* 1998;26:141–151. [PubMed: 9634136]
- Wolraich ML, Lambert W, Doffing MA, Bickman L, Simmons T, Worley K. Psychometric properties of the Vanderbilt ADHD Diagnostic Parent Rating Scale in a referred population. *Journal of Pediatric Psychology* 2003;28(8):559–567. [PubMed: 14602846]
- Zwick WR, Velicer WF. Comparison of five rules for determining the number of components to retain. *Psychological Bulletin* 1986;99(3):432–442.

Table 1

Sample Characteristics of Children With Completed SKAMP Ratings by Group

Variable	Phase 1 Representative Sample Wave 1 <i>n</i> = 1,205	Phase 2 High-Risk Group Wave 2 <i>n</i> = 190	Phase 3 High-Risk Group Wave 3 <i>n</i> = 156
	<i>N</i> (%)		
Gender			
Female	799 (66)	95 (50)	83 (53)
Male	406 (34)	95 (50)	73 (47)
Race			
African American	358 (30)	58 (31)	44 (28)
Caucasian	847 (70)	132 (69)	112 (72)
Lunch status			
Subsidized	568 (47)	100 (53)	74 (47)
Unsubsidized	637 (53)	90 (47)	82 (53)
Grade level at study entry			
Kindergarten	203 (17)	21 (11)	18 (12)
Grade 1	208 (17)	34 (18)	28 (18)
Grade 2	228 (19)	32 (17)	22 (14)
Grade 3	181 (15)	34 (18)	27 (17)
Grade 4	188 (16)	35 (18)	32 (21)
Grade 5	197 (16)	34 (18)	29 (19)
ADHD subtype ^a			
Inattentive	—	44 (23)	38 (24)
Hyperactive/impulsive	—	20 (11)	12 (8)
Combined	—	56 (29)	49 (31)
No ADHD diagnosis		70 (37)	57 (37)
Special education status at study entry			
No special services	909 (75)	128 (67)	102 (65)
Gifted	167 (14)	16 (8)	15 (10)
Emotional handicap	19 (2)	11 (6)	8 (5)
Specific learning disability	53 (4)	20 (11)	17 (11)
Other special education services	57 (5)	15 (8)	14 (9)
Discipline referrals between times 1 and 3 (<i>n</i> = 156)			
None			89 (57)
1 or 2			24 (15)
3 or more			43 (28)
		<i>M</i> (<i>SD</i>)	
Age at study entry, years	7.67 (1.77)	8.00 (1.71)	8.02 (1.72)
SES	38.96 (12.79)	37.31 (13.96)	38.00 (13.64)

Note: ADHD = attention-deficit/hyperactivity disorder; *M* = mean; *SD* = standard deviation; SES = socioeconomic status.^a diagnosis determined by DISC-IV.

Table 2
Factor Analyses Fit Indices by Subgroup Comparing One, Two, and Three Factors

	Goodness of Fit				CFI	NNFI	AIC
	χ^2	df	p	RMSEA			
One factor							
Full sample	2,643.26	35	<.00	.25	.85	.81	2,573.26
Two factors							
Full sample	1,716.54	34	<.00	.20	.90	.87	1,648.54
Male	717.36	34	<.00	.22	.88	.84	649.36
Female	1,094.25	34	<.00	.20	.90	.87	1,026.25
African American	728.34	34	<.00	.24	.87	.82	660.34
Caucasian	1,155.95	34	<.00	.20	.90	.87	1,087.95
Any concern	772.38	34	<.00	.20	.89	.86	704.38
No concern	1,304.99	34	<.00	.24	.86	.82	1,236.99
Three factors							
Full sample	1,103.77	32	<.00	.17	.94	.91	1,039.77
Male	475.49	32	<.00	.19	.92	.89	411.49
Female	757.64	32	<.00	.17	.93	.91	693.64
African American	516.53	32	<.00	.21	.91	.87	452.53
Caucasian	781.09	32	<.00	.17	.93	.90	717.09
Any concern	900.24	32	<.00	.19	.90	.87	836.24
No concern	494.16	32	<.00	.16	.93	.91	430.16

Note: AIC = Akaike information criteria; CFI = comparative fit index; *df* = degrees of freedom; NNFI = nonnormed fit index; RMSEA = root mean square error of approximation.

Table 3
Correlations Between the SKAMP, SNAP-IV, and Measures of Impairment

Measure	SKAMP Total			Attention		Department	
	n	r _s	τ	r _s	τ	r _s	τ
Teacher/school							
SNAP-IV							
Overall	1,205	.93	.75	.88	.67	.86	.64
Inattention	1,205	.93	.74	.93	.74	.77	.54
Hyperactivity/impulsivity	1,197	.79	.55	.70	.47	.85	.59
Oppositional defiance	1,197	.71	.45	.62	.39	.79	.50
Disciplinary referrals	190	.46	.30	.41	.27	.45	.29
VADTRS impairment	73	.52	.39	.49	.36	.47	.36
Parent							
SNAP-IV							
Overall	1,204	.46	.31	.43	.29	.43	.28
Inattention	1,204	.49	.33	.49	.32	.41	.27
Hyperactivity/impulsivity	1,204	.42	.28	.38	.25	.42	.28
Oppositional defiance	1,204	.31	.20	.28	.18	.33	.21
Columbia Impairment Scale (CIS)							
Overall CIS score	156	.09 ^c	.06 ^c	.11 ^c	.07 ^c	.02 ^c	.02 ^c
School-related item scores							
School behavior	156	.35	.27	.31	.25	.33	.26
Schoolwork	156	.18 ^a	.14 ^a	.23 ^b	.17 ^a	.07 ^c	.05 ^c
VADPRS impairment	106	.39	.28	.39	.28	.28 ^b	.21 ^b

Note: Values are weighted for sample design and nonparticipation. r_s (rho) = Spearman's correlation coefficient. τ (tau) = Kendall's tau, a correlation coefficient. All effects for the SNAP-IV teacher and parent, disciplinary referrals, and the CIS School Behavior items were significant at the .001 level. SKAMP = Swanson, Kotkin, Agler, M-Flynn, and Pelham Scale; SNAP-IV = Swanson, Nolan, and Pelham-IV questionnaire; VADPRS = Vanderbilt ADHD Diagnostic Parent Rating Scale; VADTRS = Vanderbilt ADHD Diagnostic Teacher Rating Scale.

^aSignificant at .05.

^bSignificant at .01.

^cNot significant.

Table 4
SKAMP Subscale Scores by Gender, Race, and Level of Concern

	Attention Subscale Level of Concern				Department Subscale Level of Concern			
	Overall	No Concern	Nonspecific	ADHD	Overall	No Concern	Nonspecific	ADHD
Full sample (<i>n</i>)	1,205	644	332	229	1,205	644	332	229
<i>SD</i>	.93	.65	.94	1.06	.77	.54	.76	.97
Mean	.87	.46	1.08	1.51	.61	.32	.73	1.11
Median	.50	.17	1.00	1.50	.25	.00	.50	1.00
90th percentile	2.50	1.33	2.50	2.83	2.00	1.00	2.00	2.50
Males (<i>n</i>)	406	167	116	123	406	167	116	123
<i>SD</i>	1.20	.95	1.14	1.20	.99	.69	.95	1.09
Mean	1.13	.62	1.30	1.62	.78	.38	.90	1.19
Median	.83	.33	1.33	1.83	.50	.25	.75	1.25
90th percentile	2.67	1.83	2.67	2.83	2.25	1.25	2.25	2.50
Females (<i>n</i>)	799	477	216	106	34	61	799	477
<i>SD</i>	.71	.49	.79	.84	.60	.47	.61	.79
Mean	.64	.36	.90	1.28	.46	.28	.59	.96
Median	.33	.17	.67	1.17	.25	.00	.25	.75
90th percentile	1.83	1.00	2.17	2.83	1.50	1.00	1.75	2.25
African American (<i>n</i>)	358	178	111	69	358	178	111	69
<i>SD</i>	1.17	.92	1.05	1.20	1.00	.79	.92	1.12
Mean	1.27	.74	1.49	1.94	.92	.53	1.04	1.48
Median	1.00	.33	1.50	2.17	.75	.25	1.00	1.50
90th percentile	2.83	2.33	2.83	3.00	2.25	1.50	2.25	2.75
Caucasian (<i>n</i>)	847	466	221	160	847	466	221	160
<i>SD</i>	.72	.46	.76	.88	.58	.36	.58	.79
Mean	.61	.31	.75	1.19	.41	.20	.48	.84
Median	.17	.00	.50	1.17	.25	.00	.25	.50
90th percentile	1.83	.83	2.00	2.50	1.25	.75	1.50	2.25

Note: Weighted for sample design and nonparticipation. ADHD = attention-deficit/hyperactivity disorder; *SD* = standard deviation.

Table 5

SKAMP Subscale Scores by Race and Poverty

	N	Total Score		Attention Subscale		Depotment Subscale	
		M	SD	M	SD	M	SD
African American							
1. Poor	321	1.16 _a	1.07	1.30 _a	1.19	.95 _a	1.03
2. Nonpoor	37	.66 _b	.68	.77 _b	.80	.50 _b	.60
Caucasian							
3. Poor	247	.73 _c	.77	.83 _c	.86	.58 _c	.75
4. Nonpoor	600	.43 _b	.53	.50 _b	.63	.32 _b	.47

Note: Poor vs. non-poor established by free or reduced lunch status. In a given column, means with different subscripts differ significantly on the basis of Tukey's post-hoc comparisons following significant omnibus multivariate analysis of variance.