Cognitive Function and Behavior of Children With Adenotonsillar Hypertrophy Suspected of Having Obstructive Sleep-Disordered Breathing
Paul M. Suratt, Mario Peruggia, Lynn D’Andrea, Robert Diamond, Jeffrey T. Barth, Margarita Nikova, Vito A. Perriello, Jr and Michael L. Johnson
*Pediatrics* 2006;118:771-781; originally published online Aug 7, 2006;
DOI: 10.1542/peds.2006-0173

This information is current as of September 11, 2006

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://www.pediatrics.org/cgi/content/full/118/3/e771
Cognitive Function and Behavior of Children With Adenotonsillar Hypertrophy Suspected of Having Obstructive Sleep-Disordered Breathing

Paul M. Suratt, MD, Mario Peruggia, PhD, Lynn D'Andrea, MD, Robert Diamond, PhD, Jeffrey T. Barth, PhD, Margarita Nikova, MD, Vito A. Perriello, Jr, MD, Michael L. Johnson, PhD

Department of Internal Medicine, Pulmonary Critical Care Division, and Department of Pediatrics, University of Virginia Medical Center, Charlottesville, Virginia; Department of Statistics, Ohio State University, Columbus, Ohio; Department of Psychiatric Medicine and Department of Pharmacology, University of Virginia Medical School, Charlottesville, Virginia; Pediatric Associates of Charlottesville, Charlottesville, Virginia

The authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVE. The purpose of this study was to determine whether risks of impaired cognitive function could be predicted for children or groups of children with adenotonsillar hypertrophy who were suspected of having obstructive sleep-disordered breathing, from historical and polysomnographic variables used separately or in combination.

METHODS. We studied 114 consecutive 6- to 12-year-old children with adenotonsillar hypertrophy, who were referred because of suspected obstructive sleep-disordered breathing, with questionnaires, assessment of tonsil size, general and memory cognitive tests, and attended polysomnography with the use of nasal pressure recording to detect flow.

RESULTS. There were important significant relationships between snore group (snored every night versus less often), sleep efficiency, and race and 2 of 3 general cognitive tests (vocabulary and similarities). Significant but weaker relationships were observed between sleep latency and 2 memory indices (verbal memory and general memory) and between sleep efficiency and 2 behavior indices (attention-deficit/hyperactivity disorder summary and hyperactive-impulsive summary). The number of episodes of apnea and hypopnea per 1 hour of sleep predicted the vocabulary score as well as did the snore group, but it did not predict other tests as well as other variables. Tonsil size did not predict any cognitive or behavior score. Confidence intervals for group means were small, whereas prediction intervals for individual children were large.

CONCLUSIONS. Risk of impaired cognitive function and behavior can be predicted from snoring history, sleep efficiency, sleep latency, and race but not tonsil size. The combination of snoring history and polysomnographic variables predicted impaired cognitive scores better than did either alone. The snoring history predicted

www.pediatrics.org/cgi/doi/10.1542/peds.2006-0173
doi:10.1542/peds.2006-0173

Key Words
sleep, obstructive sleep-disordered breathing, cognitive function, adenotonsillar hypertrophy

Abbreviations
OSDB—obstructive sleep-disordered breathing
ADHD—attention-deficit/hyperactivity disorder
AHI—apnea/hypopnea index
CI—confidence interval

Accepted for publication Mar 31, 2006
Address correspondence to Paul M. Suratt, MD, Box 800546, University of Virginia Medical Center, Charlottesville, VA 22908. E-mail: ps4p@virginia.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2006 by the American Academy of Pediatrics
more test scores than the number of episodes of apnea and hypopnea per 1 hour of sleep.

Children with habitual snoring or obstructive sleep-disordered breathing (OSDB) are reported to have impaired academic and cognitive performance.\(^1\) It is not clear whether all such children are at risk of having impaired performance or whether there are markers of OSDB and sleep that predict higher risk for individual children or for groups of children. Most studies of cognitive function in children with adenotonsillar hypertrophy that used polysomnography and cognitive tests with well-established normative results have been relatively small, involving <50 children. One of the few larger studies using these methods found significant relationships between several cognitive tests and a single sleep or breathing variable, with 1-tailed tests of Pearson correlation coefficients.\(^2\) Another such study found that the number of arousals per hour was related to several cognitive scores.\(^3\) These important studies selected children from school populations; they did not focus on children with adenotonsillar hypertrophy and did not assess tonsil size.

We are not aware of large studies that present graphical displays of the data or confidence intervals (CIs) for the derived predictive relationships, which would make it possible for readers to assess the precision of such relationships. In addition, we are not aware of large studies that describe whether combinations of sleep and breathing variables might be more helpful in predicting cognitive function than individual variables. To our knowledge, all previous studies of cognitive function in children with OSDB used nasal thermistors to detect apnea and hypopnea. This technique was shown to miss many episodes of apnea and hypopnea detected by nasal pressure signals.\(^4\)\(^-\)\(^9\) Nasal pressure recording is now the method preferred by the American Academy of Sleep Medicine Task Force.\(^10\) It has therefore been suggested that correlations between respiratory events and morbidity might improve with the use of nasal pressure recordings.\(^11\) The overall purpose of this study was to determine whether it is possible to predict the risk of impaired cognitive scores and behavior in children or groups of children with adenotonsillar hypertrophy suspected of having OSDB from historical and polysomnographic variables used separately or in combination.

**METHODS**

These studies were approved by the University of Virginia Human Investigation Committee. All participants were volunteers who signed informed assent or consent forms and whose parents signed informed consent forms for the study. We recruited 116 children, ranging in age from 6 to 12 years. These children were consecutive participants in a National Institutes of Health-sponsored study on the effect of OSDB attributable to adenotonsillar hypertrophy on behavior, cognitive performance, and growth. Because subjects in this larger study needed to remain still for 5 minutes in an MRI scanner, subjects <6 years of age were not included in the study. All children with adenotonsillar hypertrophy who were thought to have OSDB and were examined in the authors’ clinics were offered participation in this study. Children were excluded if they had a pulmonary, cardiac, neurologic, or auditory disorder or were receiving any medications, including those for attention-deficit/hyperactivity disorder (ADHD).

A sleep questionnaire was administered to all parents, to solicit information about snoring and breathing during sleep.\(^12\) Parents were asked to describe the frequency of their children’s snoring as (1) never, (2) rarely (less than once per month), (3) occasionally (1–4 times per month), (4) frequently (more than once per week), or (5) most nights. They were also asked how loud the snoring was, (1) does not snore; (2) faint, cannot hear unless near the child; (3) light, can hear in the same room but not too loud; (4) moderate, easy to hear; or (5) very loud, bothersome and cannot ignore. Parents were asked whether their child has difficulty or struggles to breathe during sleep, whether the chest “caves in” or “see-saws” during sleep, and whether the child stops breathing during sleep, falls asleep in school, naps after school, or complains of feeling tired, all using the same scale as used for snoring. They were also asked whether the child was a mouth-breather never, sometimes, or always and whether someone else in the family had sleep apnea. Parents also responded to a questionnaire about the child’s health, including whether they had ear infections and what medications they were taking.

Subjects were studied in the University of Virginia General Clinical Research Center. Tonsil size was rated on examination as 1 to 4+, as described by Brodsky.\(^13\) At 1:30 PM, children underwent cognitive function studies; these consisted of 3 subtests of the Wechsler Intelligence Scale for Children, Third Edition,\(^14\) namely, vocabulary, similarities, and block design. A difference of 3 points on each of these tests represents 1 SD. Children also performed the Wide Range Assessment of Memory and Learning test,\(^15\) which has summary scores for each of the following indices: verbal memory, visual memory, learning, and an overall summary, the general memory index. A difference of 15 points on each of these tests represents 1 SD. We also administered the Connors Continuous Performance Test,\(^16\) a computer test designed to measure sustained attention. Testing was performed by 2 experienced neuropsychology technicians. The parent accompanying the child to the General Clinical Research Center completed the Connors Parent Behavior Rating Scale, long form,\(^17\) which is used to detect behavioral problems in children and has scores for 7 factors, including oppositional, cognitive problems/inattention, hyper-
activity, anxious/shy, perfectionism, social problems, and psychosomatic problems, as well as summary scores for ADHD, inattentive, and hyperactive-impulsive.

In the evening, subjects underwent overnight polysomnography with a Sandman sleep system (Sandman Sleep Diagnostics, Kanata, Ontario, Canada), with conventional techniques including electroencephalography, electro-oculography, submental electromyography, nasal airflow measurement with a nasal cannula attached to a pressure transducer (DP 45; Validyne Engineering Inc, Northridge CA), oral airflow measurement with a thermistor, pulse oximetry, and chest and abdominal movement detection with respiratory inductance plethysmography (NonInvasive Monitoring Systems, North Bay Village, FL). Parents and children were allowed to select a lights-out time for the study that approximated their bedtime at home (but not after 10 PM). One sleep technician, who was blinded to the results of the questionnaire and all other studies, analyzed the studies by scoring sleep with conventional methods for children and scoring episodes of apnea, hypopnea, and oxyhemoglobin desaturation of >3%. All studies were reviewed by Dr D’Andrea or Dr Suratt, who were blinded to results of the questionnaire and cognitive studies. Episodes of apnea were characterized by reductions in flow to <20% of normal for ≥6 seconds or 2 breaths and episodes of hypopnea by reductions in flow to <60% of normal for ≥2 seconds or 2 breaths. Apnea and hypopnea were reported to be obstructive when the chest and abdomen moved and central when chest and abdominal movements were absent. They were quantitated with an apnea/hypopnea index (AHI), which is the number of episodes of apnea and hypopnea per 1 hour of sleep. Because criteria for arousals have not yet been developed for children, arousals were defined as recommended in the American Sleep Disorders Association Task Force report and included respiratory-related, technicain-induced, and spontaneous arousals. Arousals were expressed as total arousals per 1 hour of sleep. Sleep latency was the first of 3 consecutive epochs of stage 1 sleep, the first epoch of stage 1 sleep followed by any other stage of sleep, or the first epoch of any other stage of sleep. Sleep efficiency was the ratio of total sleep time to total time in bed, expressed as a percentage. Because of errors in the commercial software calculations, oxyhemoglobin saturation values were extracted from the Sandman event files into ASCII files. Using a previously described program, we calculated the lowest oxyhemoglobin saturation value during sleep. Two subjects’ files became corrupted and it was not possible to calculate their values; consequently, those subjects were not included in this analysis, leaving 114 subjects in the study.

Statistical analyses were performed with S-Plus software (Insightful Corp, Seattle, WA). Multivariate linear regression analysis was used to relate cognitive and behavloral tests to sleep and breathing parameters and potential confounders. Sleep variables included sleep latency, sleep efficiency, percentage of sleep in non–rapid eye movement and rapid eye movement sleep, and both arousals and sleep stage shifts per 1 hour of sleep. Breathing variables included sleep questionnaire data, the AHI, minimal oxyhemoglobin saturation, and percentage of time with saturation of <90%. Potential confounders included gender, age, race, health insurance status, and type of school attended (private versus public). Snoring frequency was analyzed as a dichotomous variable, with a high snoring category including all subjects whose parents described their snoring frequency as 5 and a low snoring category including all other subjects. Race was also analyzed as a dichotomous variable, with black children in one group and “other” children (white and Asian children) in the other group. Comparisons between the 2 groups were performed with the 2-sample Wilcoxon rank-sum test.

RESULTS

Study Group
We recruited 116 subjects; data from 2 subjects could not be used because of corrupted files, as described above, leaving 114 subjects in the study. There were 58 boys and 56 girls; 34 children were black and 80 were other (78 white and 2 Asian). The mean age was 8.5 ± 1.9 years. Tonsil size on examination was 1+ for 1 subject, 2+ for 22 subjects, 3+ for 60 subjects, and 4+ for 31 subjects. The mean BMI z score was 1.07 ± 1.30 (range: −2.93 to 3.17). Quartile boundaries for the distribution of AHI values were 0, 1.350, 2.615, 6.69, and 37.79, which indicates that, for example, 25% of subjects had AHI values between 0 and 1.350 and 25% between 1.350 and 2.615. The median AHI was 2.615, whereas mean AHI was 5.58. Graphical representations of AHI and other sleep and breathing variables are shown in Figs 1 to 3. A summary of all significant findings is displayed in Tables 1 and 2.

Wechsler Tests
The cognitive tests with the strongest relationships to sleep and breathing variables were 2 of the Wechsler subtests, namely, vocabulary and similarities. For vocabulary, the most important of these variables were snoring group and sleep efficiency. The most important confounder for vocabulary was race. None of the other potential confounders (gender, age, health insurance status, and type of school attended) seemed to be important.

Figure 2 presents scatter plots of the vocabulary scores versus sleep efficiency and illustrates several important relationships. First, there was some evidence of a decreasing relationship between vocabulary score and sleep efficiency. Second, the vocabulary scores were
higher in the low snoring group than in the high snoring group (median scores of 12 for the low snoring group and 10 for the high snoring group; P = .0047) (Table 1). Third, the vocabulary scores were lower in the black group than in the other group (median scores of 9 for the black group and 11 for the other group; P = .003) (Table 1).

The basic features revealed by Fig 2 are incorporated in the following well-fitting linear regression model (model 1) for the prediction of vocabulary scores: vocabulary = 17.23 – (2.37 × snore group) + (1.593 × race group) – (0.073 × sleep efficiency) \( (r^2 = 20.09\% ; P < .0001\) for model utility test). The term snore group is the dichotomous indicator of snoring (0 for subjects in the low snoring group and 1 for subjects in the high snoring group). Its negative coefficient of \(-2.372\) \( (P = .0007)\) reflects the negative impact of high snoring levels on predicted vocabulary scores. The term race is an indicator of race (0 for subjects in the black group and 1 for subjects in the other group). Its positive coefficient of 1.59 \( (P = .0110)\) constitutes an adjustment for the difference in vocabulary scores between the 2 groups. After accounting for the race and snoring terms, the term sleep efficiency has a negative coefficient of \(-0.07265\) \( (P = .0222)\), providing an indication that, within a given race and snoring group, lower vocabulary scores should be predicted for children with higher sleep efficiencies (higher sleep efficiencies might occur among sleep-deprived children).

Figure 3 presents scatter plots of the vocabulary scores versus \(\log_{10}(AHI + 1)\). We used a logarithmic transformation of AHI because the relationship with the vocabulary scores seemed to be more nearly linear on this scale. Figure 3 provides evidence of a moderate decreasing trend in vocabulary scores as \(\log_{10}(AHI + 1)\) increased. Adding \(\log_{10}(AHI + 1)\) as a predictor to model 1 did not improve the model substantially. The multiple correlation coefficient of model 1 including \(\log_{10}(AHI + 1)\) increased from 20.09% to 21.53%, but the P value for the coefficient of \(\log_{10}(AHI + 1)\) was not significant \( (P = .1596)\). However, the following model (model 2), which includes \(\log_{10}(AHI + 1)\) as a predictor in place of snore group, has qualitative features similar to those of model 1: vocabulary = 16.50 – \([1.95 \times \log_{10}(AHI + 1)] + [1.63 \times (race group)] + [0.071 \times sleep efficiency] (r^2 = 16.65\%\% ; P = .0002 for model utility test). The P value for coefficient \(\log_{10}(AHI + 1)\) was .0085, the P value for race group was .0113, and the P value for sleep efficiency was .0274. The overall strength of the linear dependence of the vocabulary scores on the new set of predictors similar to that of model 1, making the 2 models essentially equivalent. This is related to the finding that subjects in the high snoring group had higher values for \(\log_{10}(AHI + 1)\) than did subjects in the low snoring group (median scores of 0.66 \([AHI = 3.54]\) in the high snoring group and 0.28 \([AHI = 0.92]\) in the low snoring group; \( P < .0001\)). There was also weak evidence of a decreasing trend in vocabulary scores as \(\log_{10}(arousals per hour + 1)\) increased \( (r^2 = 4.50\% ; P = .0169)\).

Figure 4 presents scatter plots of similarities scores versus sleep efficiency and illustrates several important relationships. First, there is some evidence of a decreasing relationship between similarities score and sleep efficiency. Second, the similarities scores were higher in the low snoring group than in the high snoring group (median scores of 13 in the low snoring group and 10 in the high snoring group; \( P = .0062\)) (Table 1). Third, the similarities scores were lower in the black group than in the other group (median scores of 10 in the black group and 12 in the other group; \( P = .0008\)) (Table 1).

The impact of snoring group and sleep efficiency on the similarities score is qualitatively similar to that on the vocabulary score. A well-fitting linear regression model (model 3) for the prediction of similarities scores is as follows: similarities = 17.06 – \([1.76 \times snore group] + (1.98 \times race group)] – (0.0744 \times sleep efficiency) (r^2 = 20.65\% ; P < .0001 for model utility test). The P value for
snore group was .0081, the \( P \) value for sleep efficiency was .0151, and the \( P \) value for race was .0011. The addition of \( \log_{10}(\text{AHI} + 1) \), \( \log_{10}(\text{arousals per hour} + 1) \), or sleep latency as predictors to model 3 did not produce a substantial improvement in the proportion of variation explained by the models, and none of the coefficients for these 3 variables were significant in the multivariate regression. In addition, sleep latency showed a very weak positive linear association with similarities (\( r^2 = 3.51\% \); \( P = .0458 \)). Block design (visuospatial problem-solving) scores were not well predicted by any linear relationship involving sleep or breathing variables.

**Memory Tests**

There were significant but rather weak relationships between sleep latency and both the verbal memory index and the general memory index, which is the overall memory summary score. These relationships are summarized by the following linear regression models: model 4: verbal memory index = 89.18 + (0.11 \times sleep latency) (\( r^2 = 11.51\% \); \( P = .0002 \)); model 5: general memory index = 94.57 + (0.10 \times sleep latency) (\( r^2 = 7.34\% \); \( P = .0036 \)).

With respect to model 4, there was 1 child with extreme values (sleep latency of 262.08 minutes and verbal memory index of 106). Removing this case from the analysis had a modest impact on the fit of model 4 (\( r^2 = 11.98\% \); \( P = .0002 \)). With respect to model 5, there were 2 children with extreme values. One had a general memory index of 15 and a sleep latency of 47.2 minutes, and the other has a general memory index of 124 and a sleep latency of 262.08 minutes. Removing both of these cases from the analysis had a modest impact on the fit of model 5 (\( r^2 = 6.81\% \); \( P = .0054 \)).
Behavior Tests

Children in the high snoring group had higher and more-abnormal median psychosomatic scores than did those in the low snoring group (69 vs 53 for 99 children with available scores; 2-sided $P = .0001$) (Table 1). There were also very weak relationships between both the

---

**TABLE 1** Differences in Cognitive and Behavioral Scores Within Snoring and Racial Groups, Not Adjusted for Other Variables

<table>
<thead>
<tr>
<th>Test</th>
<th>Test Score</th>
<th>Snoring Group</th>
<th>Race</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Low</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td></td>
<td>12</td>
<td>10</td>
<td>.0047</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>12.6</td>
<td>10.09</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td>3.85</td>
<td>2.89</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td></td>
<td>4–19</td>
<td>3–17</td>
<td></td>
</tr>
<tr>
<td>Similarities</td>
<td></td>
<td>13</td>
<td>10</td>
<td>.0062</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>12.6</td>
<td>10.65</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td>3.2</td>
<td>3.04</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td></td>
<td>6–19</td>
<td>1–19</td>
<td></td>
</tr>
<tr>
<td>Psychosomatic</td>
<td></td>
<td>53</td>
<td>69</td>
<td>.0001</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>54.84</td>
<td>67.23</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td></td>
<td>10.72</td>
<td>13.71</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td></td>
<td>43–90</td>
<td>43–90</td>
<td></td>
</tr>
</tbody>
</table>

---

**TABLE 2** Significant Cognitive and Behavioral Correlations

<table>
<thead>
<tr>
<th>Test Domain</th>
<th>Neurocognitive Test</th>
<th>Model</th>
<th>Significant Variables</th>
<th>$r^2$</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>General (WISC-III) Vocabulary</td>
<td>1</td>
<td>Snore group, sleep efficiency, race</td>
<td>20.09</td>
<td>.000172</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Log$_10$(AHI + 1), sleep efficiency, race</td>
<td>16.65</td>
<td>.0001599</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Snore group, sleep efficiency, race</td>
<td>20.65</td>
<td>.0000118</td>
<td></td>
</tr>
<tr>
<td>Similarities</td>
<td>4</td>
<td>Sleep latency</td>
<td>3.515</td>
<td>.04578</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Sleep efficiency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Sleep latency</td>
<td>11.51</td>
<td>.000222</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>Sleep latency</td>
<td>7.339</td>
<td>.003554</td>
<td></td>
</tr>
<tr>
<td>Memory Verbal memory index</td>
<td>4</td>
<td>Sleep efficiency</td>
<td>4.364</td>
<td>.03799</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Sleep efficiency</td>
<td>4.443</td>
<td>.0346</td>
<td></td>
</tr>
<tr>
<td>Behavior ADHD summary</td>
<td>6</td>
<td>Sleep efficiency</td>
<td>4.364</td>
<td>.03799</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>Sleep efficiency</td>
<td>4.443</td>
<td>.0346</td>
<td></td>
</tr>
<tr>
<td>Continuous performance Block design</td>
<td>None significant</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**FIGURE 4** Relationship between similarities scores and sleep efficiency. In A, subjects in the low snoring group are displayed as circles and those in the high snoring group as triangles. Similarities scores are higher in the low snoring group (see text). In B, black subjects are displayed as circles and other subjects as triangles. Similarities scores are higher in the other group (see text).
summary score for ADHD (ADHD index) and the summary score for hyperactive-impulsive (Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, symptom subscales) and sleep efficiency, based on data for 101 children with available scores, as follows: model 6: ADHD score = 32.88 + (0.33 × sleep efficiency) ($r^2$ = 4.36%; $P$ = .0380); model 7: hyperactive-impulsive score = 24.65 + (0.432 × sleep efficiency) ($r^2$ = 5.75%; $P$ = .0168). Finally, we were not able to identify any important linear relationships between sleep and breathing variables and any scores from the Connors Continuous Performance Test.

Sleep and Breathing Variables That Did Not Improve Predictions of Cognitive and Behavioral Scores

The addition of historical questionnaire data other than snoring frequency did not improve predictions of cognitive or behavioral scores. Similarly, the lowest oxyhemoglobin saturation value, the percentage of time with saturation of <90%, and tonsil size were found to be of no importance in this population for prediction of any cognitive or behavioral scores. There was also no linear relationship between BMI $z$ score or percentile score and any cognitive or behavioral test result.

Relationships Among Sleep and Breathing Variables

We observed the following relationships between sleep and breathing variables that we used to analyze cognitive function. The strongest relationship was a negative relationship between sleep latency and sleep efficiency ($r^2 = 51.32% ; P < .0001$), such that shorter sleep latencies were associated with higher sleep efficiencies. A positive relationship was found between log$_{10}$(AHI + 1) and log$_{10}$(arousals per hour + 1) ($r^2 = 29.34% ; P < .0001$). There was no significant relationship between sleep latency and log$_{10}$(AHI + 1), snore group, or arousals per hour. There was no significant relationship between sleep efficiency and log$_{10}$(AHI + 1), snore group, or arousals per hour.

Regarding oxyhemoglobin saturation, 2 children had oxyhemoglobin saturation nadirs of <67% and their AHI values all exceeded 24 (Fig 1). The most extreme case was a child with an oxyhemoglobin saturation nadir equal to 57.4% and an AHI equal to 37.8. Including these high-leverage cases in a linear regression analysis relating AHI to oxyhemoglobin nadirs would inflate unduly the importance of a linear relationship that is not supported by the bulk of the data. All remaining children had oxyhemoglobin saturation nadir values of >73% and AHI values of <25. After removal of the 2 influential cases, there was only a very small negative linear association between log$_{10}$(AHI + 1) and oxyhemoglobin saturation nadirs during sleep ($r^2 = 4.05% ; P = .0333$). The BMI $z$ score was not related to snore group, sleep latency, sleep efficiency, log$_{10}$(AHI + 1), arousals per hour, lowest oxyhemoglobin saturation value, or percentage of time with saturation of <90%.

**DISCUSSION**

In this study, children with adenotonsillar hypertrophy who were suspected of having OSDB were more likely to have impaired cognitive function if they had a history of nightly snoring and if polysomnographic studies suggested that they were sleepy. The mean degree of impairment in these tests attributable to OSDB and sleepiness is profound and similar in magnitude to the effects of lead exposure in children. For example, for non-black children, model 1 estimated that children who do not snore every night and who have a sleep efficiency of 70% during attended polysomnography would have a mean vocabulary score of 13.7 (95% CI: 12.4–15.1). Children who snore every night and have a sleep efficiency of 90% would have a mean vocabulary score of only 9.9 (95% CI: 8.9–10.9). This represents a reduction in vocabulary scores equal to 1.3 SD of the distribution of scores in the general population. Similarly, model 2 estimates that nonblack children with a sleep efficiency of 70% and an AHI of 0 would have a mean vocabulary score of 13.1 (95% CI: 11.8–14.4), whereas children with a sleep efficiency of 90% and an AHI of 10 would have a mean vocabulary score of 9.7 (95% CI: 8.5–10.8), a reduction of 1.1 SD. For Wechsler tests, a difference of 0.5 or 1 SD in scores from the population mean is considered clinically important. A change of 1 SD is equivalent to a 15-point change on the full-scale Wechsler IQ scale. For example, studies of lead exposure in children reported that children with a lifetime average blood lead concentration of up to 10 μg/dL had a 7.4-point reduction in IQ, a decline thought to be important. It is important to note, however, that the resulting prediction intervals are much wider when the models are used to predict vocabulary scores for individual children with given characteristics. For example, for the cases considered above, the 95% CIs resulting from model 1 would be much larger (7.7–19.8 and 4.0–15.9). The 95% prediction intervals resulting from model 2 would also be much larger (7.0–19.3 and 3.5–15.8). Because of the considerable overlap of the intervals, it is not possible to predict reliably the effect of OSDB and sleepiness on cognitive function for individual subjects. There are, of course, many factors that contribute to cognitive function in children, of which OSDB and sleepiness are but 2 and many of which are unaccountable. It is therefore not surprising that the data exhibit large subject-to-subject variability.

Similar findings were noted for the similarities score with model 3, showing that snoring group and sleep efficiency were related to the score in a manner comparable to their relationships with the vocabulary score. Again, CIs for mean similarities values were small but...
prediction intervals for individual subjects were very wide. Similarly, sleep latency predicted verbal memory and the memory summary score (general memory index), although these relationships were weaker than those involving the vocabulary and similarities scores.

The vocabulary subtest is viewed as the best single predictor of overall Wechsler IQ score and general cognitive functioning and is considered a strong predictor of academic success.\(^{29}\) The similarities subtest is a good indicator of verbal abstract reasoning, which is also important in new problem-solving and learning. The declines in mean vocabulary and similarities scores noted above represent losses that could put the subpopulation of children with OSBD and sleepiness at a serious disadvantage, in terms of scholastic performance. The skills measured with the verbal memory test, including retention and retrieval of contextual information presented in a listening format, are vital components of the format of instruction in schools. The deficits noted in these results are likely to be longstanding and to interact with other deficits and thus to contribute to a downward academic spiral. The vocabulary, similarities, and verbal memory tests are relatively nonengaging tasks that require a combination of listening, extended cognitive effort without environmental stimulation, and autonomous sequencing of information. This is in contrast to the other tests in our battery, which include substantial components of visual and manipulative stimuli and cuing materials. The vocabulary subtest presents the child with a single word and then requires the formulation and expression of an acceptable definition. The similarities test involves listening to pairs of words and then responding regarding how they are alike. The verbal memory subtests (story memory, sentence memory, and number/letter memory) all require the child to sit passively, to screen out distractions, and to maintain sustained attention in listening to extended presentations of information and then to recall and to retrieve that information in context or in sequence. Therefore, it seems that children with OSBD in this study have particular difficulty when confronted with nonengaging complex tasks requiring autonomous sequencing. Because success at school requires ability with these types of tasks, the difficulty children with OSBD have with these tasks may explain their poor performance at school.

Among behavioral scores, the ADHD and hyperactive-impulsive summary scores were related significantly but weakly to sleep efficiency. There was also a markedly significant difference in median parent-rated psychosomatic test scores in the 2 snoring groups: higher sleep efficiencies (more sleepy children) or more frequent snoring predicted more behavioral problems. Several other studies of children with OSBD or chronic snoring also found that these children were more hyperactive or inattentive than normal children.\(^ {11,26-28}\) Our study might have underestimated the degree of behavior problems, because we excluded children taking stimulants.

The snoring group predicted more cognitive test scores than did the AHI, despite our use of nasal pressure to detect nasal flow. The AHI may not detect episodes of OSDB in children for several reasons. First, children with OSBD often have episodes of increased upper airway resistance, followed by arousal, that are reported to be apparent only when children are studied with an esophageal catheter, which reflects respiratory effort.\(^ {29}\) Second, children with OSBD frequently have long periods of absent nasal flow, when it is difficult to know whether they are breathing adequately through their mouths. Recently, we reported that children with OSBD frequently make high-pitched inspiratory sounds during sleep, which sound as though the children are struggling to inhale through a narrow airway.\(^ {30,31}\) These events are not usually counted as episodes of apnea or hypopnea because they last for long periods of up to 300 consecutive breaths. Apnea and hypopnea also might be missed when a thermistor is used to detect airflow, although this was not the case in the current study, because we detected nasal flow with a pressure transducer.

Polysomnographic variables reflecting the propensity to go to sleep (sleep latency) and remain asleep (sleep efficiency) added significantly to our ability to predict cognitive and behavior outcomes. Although it may seem paradoxical that children with short sleep latencies and high sleep efficiencies (ie, those who sleep well in the sleep laboratory) performed worse on cognitive tests, there are several plausible explanations. First, OSBD could be disturbing sleep. O’Brien et al\(^ {6}\) found that children with OSBD had more total arousals per hour than did control subjects. Our findings were similar, in that children with higher AHI values had more total arousals per hour. Although studies using conventional methods of scoring sleep have not found disturbed sleep architecture among children with OSBD, conventional methods of scoring sleep can be insensitive to short periods of wakefulness. For example, if a child is awake for 14 seconds and asleep for 16 seconds during a 30-second epoch, then the entire 30-second epoch is scored as sleep. When sleep was studied with computerized analysis of electroencephalographic signals, several studies found that sleep was abnormal for children with OSBD. Bandla and Gozal\(^ {14}\) found that children with OSBD had less delta sleep power, reflecting slow-wave or deep sleep, and Chervin et al\(^ {12}\) reported that they had more microarousals than children without OSBD. Second, children may not be getting enough sleep at home and are sleep deprived, as suggested by Blunden et al.\(^ {33}\) Sleep-deprived children would fall asleep faster and stay asleep longer in the sleep laboratory, compared with children who are not sleep deprived. This explanation is supported by several studies indicating that many schoolchildren are getting to bed too late or getting up...
too early to get enough sleep.\textsuperscript{34,35} Although we attempted to approximate the child’s normal bedtime in the sleep laboratory, we do not know how successful we were. It is possible that some children went to bed earlier in the sleep laboratory than they do at home, which could have contributed to longer sleep latencies and lower sleep efficiencies. This possibility will need to be tested in future studies.

There was no relationship between oxyhemoglobin saturation and neurocognitive and behavioral function in our population. Urschitz et al\textsuperscript{36} reported that school performance was impaired for snoring children, without an independent effect of intermittent hypoxia (<90% saturation). Although children who snore and have significant oxyhemoglobin desaturation during sleep are clearly at risk of poor performance,\textsuperscript{5} both our study and that by Urschitz et al\textsuperscript{36} indicate that children can have poor performance related to OSDB without having oxyhemoglobin desaturation. Because our study had few subjects with severe oxyhemoglobin desaturation, it has no power to detect the existence of an association between more-severe oxyhemoglobin desaturation and cognitive impairment.

We found that race was a significant confounder in the prediction of cognitive performance, a fact well documented among Virginian children.\textsuperscript{37} Disparities in scoring on scholastic and cognitive achievement tests across racial groups are typically attributable to multiple socioeconomic factors (eg, parents’ education, family income, and intact families), as shown by Chervin et al\textsuperscript{37} in their study of children with sleep-disordered breathing. The cognitive tests we used made no adjustment for race or other socioeconomic factors. When children taking these cognitive tests are stratified according to these multiple socioeconomic factors,\textsuperscript{14} racial differences in test performance are reported to become insignificant. Our study sample, however, was not stratified to equalize socioeconomic factors across race. Therefore, it became important to control for socioeconomic conditions in the analysis by including an indicator of race. OSDB is reported to occur more frequently in black children.\textsuperscript{14} Therefore, OSDB can have a particularly devastating effect on cognitive function in populations of black children, who might already be at a disadvantage because of their socioeconomic conditions.

Our findings of abnormal vocabulary and verbal memory in children with OSDB are similar to those of the other large study of children with polysomnography, the Tucson study.\textsuperscript{2} That school cohort study included 149 children who had returned a questionnaire about breathing during sleep (20% of questionnaires were returned) and whose parents agreed to participate in the study. Children were studied with unattended home polysomnography with nasal thermistors. The study found that subjects with AHI values of ≥5 had impairments in auditory verbal learning and delayed recall, compared with children with AHI values of <5 (P = .01 to P = .04). The authors found no differences in verbal or full-scale IQ, reading, math, or learning. By using 1-tailed Person correlation tests, they found that the degree of OSDB was correlated with full-scale and performance IQ scores and immediate recall and applied problems cognitive tests and that the percentage of stage 1 sleep was correlated with 6 of 9 verbal learning tests (all correlations were significant at the P < .05 level). Therefore, the study showed small but significant changes in several verbal learning tests. That study is complementary to ours, in that it might better address the effect of OSDB in the general population; however, our study, which was limited to children with adenotonsillar hypertrophy who were suspected of having OSDB, might better address the issue of the impact of OSBD on children with adenotonsillar hypertrophy who seek medical help.

There are both advantages and disadvantages to our study design. By recruiting children with suspected OSDB, our sample reflects the type of children likely to be seen by physicians for evaluation of OSDB. To our knowledge, this study represents the largest sample of children with adenotonsillar hypertrophy who were suspected of having OSDB who were studied in a sleep laboratory with polysomnography and with cognitive and behavioral tests that have well-standardized normative values. A disadvantage of this design, however, is that our sample may not be typical of the general population, because subjects who were performing poorly at school might have been more likely to be included in our study. An advantage of our methods of analysis is that we made no assumptions about the definition of OSDB, because we did not define cutoff values of AHI, oxyhemoglobin saturation, and arousals that separated normal children from children with OSDB. Because there are no generally accepted empirical data to support a definition of OSDB in children,\textsuperscript{4} avoiding such cutoff levels allowed us to consider OSDB as a continuum rather than as being either present or absent. We did not have to combine children with different degrees of severity in the same group for statistical analysis. This allowed us flexibility and realism in modeling the relationships between sleep and breathing variables and cognitive function. Also, we did not have to create a definition of OSDB that included an amalgam of several measurements, such as snoring history and polysomnographic measures. A disadvantage of our study design is that, by selecting for children with adenotonsillar hypertrophy who were suspected of having OSDB, we had many children who snored every night and fewer children in the less-frequent snoring categories. The split we propose (a high snoring category including all subjects whose parents described their snoring frequency as 5 and a low snoring category including all other subjects) shares the limitations of all dichotomizations based on arbitrary cutoff values but
has the advantage of resulting in 2 fairly balanced groups. On the basis of our analysis, we can state that children who snore every night are at increased risk of cognitive impairment, but we cannot establish finer risk differentiations for the other groups.

This is the largest attended polysomnographic study of children with adenotonsillar hypertrophy and, to our knowledge, the only one to use nasal pressure cannulae to detect apnea and hypopnea. The advantage of attended polysomnography is that disconnected or faulty probes can be replaced during the study, whereas in unattended studies the study must be scored without the inoperative channel or scored for a shorter time period, omitting the period when the signal is inoperative. The disadvantage of attended studies is that subjects may not sleep as well as at home.

CONCLUSIONS
This study shows that, for children with adenotonsillar hypertrophy who are suspected of having OSDB, nightly snoring, higher AHI values, and high sleep efficiencies or short sleep latencies (in the sleep laboratory) predict impaired cognitive performance, especially with general verbal, abstract reasoning, and verbal memory tasks. The degree of impairment of intellectual function in particular was profound and similar to that of lead exposure. A history of snoring and polysomnography were complementary in predicting performance; predictions using both were better than predictions using either alone. The use of nasal pressure recordings did not enable the AHI to supplant a history of snoring. Tonsil size did not predict any cognitive or behavioral test score. Although the findings were highly significant for the group as a whole, it is not possible to produce accurate predictions of cognitive function for individual subjects based on sleep and breathing variables, because of the large subject-to-subject variability reflected in the width of the CIs.

ACKNOWLEDGMENTS
This work was supported by the National Heart, Lung, and Blood Institute (grant HL62401, to P.M.S.) and the University of Virginia General Clinical Research Center (grant M01 RR00847).

We acknowledge Tracy Puffenbarger for expert assistance with the polysomnography studies, Nancy Mays and Pam McArdle for performing the neuropsychological tests, and Drs Charles M. Johnson III and Stephen Early for subject referrals.

REFERENCES
blood lead concentrations below 10 micrograms per deciliter. 


