Medications Do Not Necessarily Normalize Cognition in ADHD Patients
C. Thomas Gualtieri and Lynda G. Johnson
J Atten Disord 2008; 11; 459 originally published online Oct 12, 2007;
DOI: 10.1177/1087054707305314

The online version of this article can be found at:
http://jad.sagepub.com/cgi/content/abstract/11/4/459

Additional services and information for Journal of Attention Disorders can be found at:
Email Alerts: http://jad.sagepub.com/cgi/alerts
Subscriptions: http://jad.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav
Citations http://jad.sagepub.com/cgi/content/refs/11/4/459
Medications Do Not Necessarily Normalize Cognition in ADHD Patients

C. Thomas Gualtieri
Lynda G. Johnson
North Carolina Neuropsychiatry Clinics, Chapel Hill and Charlotte

Objective: Although ADHD medications are effective for the behavioral components of the disorder, little information exists concerning their effects on cognition, especially in community samples. Method: A cross-sectional study of ADHD patients treated with three different ADHD drugs was conducted. Patients’ performance on a computerized neurocognitive screening battery was compared to untreated ADHD patients and normal controls (NML). A total of 177 ADHD patients aged 10 to 18, achieved a favorable response to one of the following medications: Adderall XR (AMP), atomoxetine (ATMX), and Concerta (MPH-OROS) compared to 95 untreated ADHD patients and 101 NML. Results: Significant differences were detected between normals and untreated ADHD patients. Treated patients performed better than untreated patients but remained significantly impaired compared to normal subjects. Conclusion: Even with optimal treatment, based on parents’ and teachers’ opinions, subtle and not-so-subtle neurocognitive impairments persisted in the ADHD patients. Some ADHD patients may require additional educational assistance, even in the face of successful medication treatment. (J. of Att. Dis. 2008; 11(4) 459-469)

Keywords: ADHD; cognition; amphetamine; methylphenidate; atomoxetine

Psychostimulant drugs affect diverse cognitive functions, including executive function, reaction time, fine motor coordination, and various aspects of attention. Neuropsychological studies of ADHD children and adults indicate impairments in many cognitive areas: selective attention (Brodeur & Pond, 2001); memory (Muir-Broaddus, Rosenstein, Medina, & Soderberg, 2002; Roth et al., 2004); reaction time (Leth-Steensen, Elbaz, & Douglas, 2000) and information processing speed (Weiler, Bernstein, Bellinger, & Waber, 2000); motor speed (Mitchell, Chavez, Baker, Guzman, & Azen, 1990) and visuomotor ability (Kalff et al., 2002); and executive control functions, such as set-shifting (Cepeda, Cepeda, & Kramer, 2000), inhibitory control (Schachar et al., 2002), and working memory (Barnett et al., 2001). It has never been established, however, that the cognitive effects of stimulant drugs are central to their therapeutic utility. In ADHD treatment, cognitive and behavioral improvement are not necessarily dissociated, but neither are they closely correlated (Konrad, Gunther, Hanisch, & Herpertz-Dahlmann, 2004).

It has been suggested that neurocognitive testing has predictive value in determining individual differences in drug response (Mehta, Goodyer, & Sahakian, 2004), although studies have also shown that cognition may improve in ADHD patients, excluding behavioral improvement (Gimpel et al., 2005; Vance, Maruff, & Barnett, 2003). Conversely, behavior may improve absent detectable changes in cognition (Everett, Thomas, Cote, Levesque, & Michaud, 1991; Lufi, Parish-Plass, & Gai, 1997). Computerized tests of attention and vigilance

Authors’ Note: Drs Gualtieri and Johnson are two of the developers of the CNS Vital Signs screening battery. Dr Gualtieri has conducted clinical trials on behalf of Astra-Zeneca, Bristol-Myers Squibb, Celltech, Cephalon, Eli Lilly, Glaxo-Smith-Kline, Medeva, Organon, Shire, Wyeth-Ayerst, and UCB. He has been a speaker for and/or consultant to Eli Lilly, GSK, Pfizer, Shire, and Wyeth. This research was supported by North Carolina Neuropsychiatry, PA, in Chapel Hill and Charlotte. No external support was sought or received on behalf of this research. Address correspondence to Dr C. Thomas Gualtieri, NC Neuropsychiatry, 400 Franklin Square, 1829 East Franklin Street, Chapel Hill, NC 27514; e-mail: tg@ncneuropsych.com.
(Continuous Performance Tests [CPTs]), widely used in the evaluation and treatment of ADHD, have been criticized for their low specificity and sensitivity (Lovejoy & Rasmussen, 1990; Trommer, Hoeppner, Lober, & Armstrong, 1988), and the correspondence between impulsive errors on CPTs and behavioral impulsivity has not been established (Abikoff & Klein, 1992). When used for assessment of medication efficacy, the applicability of results to the patient’s natural environment is unproven (Aman & Turbott, 1991; Cohen, Kelly, & Atkinson, 1989) or even absent (Elia, Borchering, Rapoport, & Keysor, 1991). CPTs are not consistently sensitive to stimulant effects (Fischer & Newby, 1998).

When ADHD studies address the issue of cognition, the methods used tend to demonstrate that treated patients perform better than untreated patients on neuropsychological tests or that patients improve on neuropsychological measures after they are treated (Aggarwal & Lillystone, 2000; Gimpel et al., 2005). Few studies, however, evaluate posttreatment performance in ADHD patients relative to normal controls (NML). In 1991, Everett et al. compared untreated ADHD children to NML on two tests of executive function, the Stroop Test (ST) and the Wisconsin Cart Sort. At baseline, the ADHD group was significantly impaired compared to normals. After a year of treatment, they improved on the Wisconsin but remained impaired on the ST (Everett et al., 1991). In 1993, Risser and Bowers compared 10 ADHD children on stimulant medication to normal children on a number of cognitive measures, including the Wechsler Developmental Index, the Bender Visual-motor Gestalt Test, and the Benton Visual Retention Test. In spite of treatment, the ADHD children had persistent neuropsychological deficits (Risser & Bowers, 1993). In another study, stimulant treatment in ADHD children led to improvement in simple but not choice reaction time (Adams, 1982).

We have found only one study that generated contrary results. Barnett et al., using a computerized test (CogState), reported that 21 stimulant-treated ADHD children performed as well as NML on a spatial working memory task, whereas 27 untreated ADHD children performed significantly worse (Barnett et al., 2001). The CogState battery, however, captures data on several other cognitive functions, and these were not reported in the article.

The question we address here is not whether cognitive and behavioral changes are correlated when ADHD patients are treated with psychostimulant medications; it is fairly well established that they are not. The question we pose is whether, in a community-treated sample of patients, successful treatment normalizes neurocognitive performance. We used a cross-sectional approach, examining the clinical status of all the ADHD patients in a particular locus at a point in time. The question we posed is, how effective are stimulant (and related) drugs as they are used by practitioners in the real world, at least with respect to neurocognition.

Methods and Materials

This was a cross-sectional descriptive study of clinic attendees, children, and adolescents to evaluate their neurocognitive performance with respect to medication treatment. The subjects had responded satisfactorily to one of the three most frequently prescribed medications for ADHD in the United States: Adderall XR (AMP, four salts of d- and l-amphetamine in an extended release preparation), atomoxetine (ATMX), and Concerta (MPH-OROS, methylphenidate isomers in an osmotic release capsule) compared to NML and to untreated patients with ADHD.

Subjects

All of the ADHD patients at two neuropsychiatry clinics (in Chapel Hill and Charlotte, NC) are administered a computerized screening battery of cognitive tests as part of their initial evaluation and at appropriate intervals during follow-up. They are requested to allow the clinics to store their data in a central research database after it is appropriately deidentified. The database includes basic demographic data, diagnoses, medication status, and test results.

The subjects of this investigation were children and adolescents with all types of ADHD. Initial diagnoses were made by treating clinicians on the basis of DSM-IV-TR criteria, parent and teacher rating scales. Diagnoses were reviewed by a second senior psychiatrist (T.G.) or neuropsychologist (L.J.) prior to inclusion in the database and then again prior to inclusion in this study. Subjects were outpatients who had been treated with one of the three medications listed above, and the dose had been titrated to an optimal clinical response. The Ss selected for this investigation had been on a stable dose of drug for at least 4 weeks.

The patients were not randomly assigned to one drug or another; rather, they themselves chose which medication to use. As a rule, newly diagnosed ADHD patients are given two prescriptions, one for AMP and one for MPH-OROS. Patients are instructed to take one drug for 2 weeks and to titrate to optimal dose, and then to try the other. The titration schedule for AMP would be 20 or 40 mg qam and for MPH-OROS 18–36–54–72 mg qam. After 4 weeks, the patients return to the clinic and a
A decision is made on clinical grounds to continue on one drug or another. After another month, they return to the clinic and are tested on what they (and their parents) consider to be the optimal dose of the preferred drug. This method accounted for the 81 subjects on AMP and the 51 Ss on MPH-OROS.

A group of patients on stable and clinically effective doses of the nonstimulant ADHD drug, ATMX, were chosen as a comparison group. The 45 patients on ATMX were children and adolescents who were taking ATMX because they had experienced side effects on psychostimulants.

NML (N = 101) were children and adolescents with no active psychiatric, neurological, or medical conditions nor any history of participating in a normative study for the CNS Vital Signs (CNSVS) database. ADHD controls (N = 95) were newly diagnosed patients who had not yet begun to take any medications. Demographic data on the five groups are presented in Table 1.

### Cognitive Evaluation

Patients’ neurocognitive performance was measured on a computerized battery of tests. The CNSVS battery contains seven tests that are widely used by neuropsychologists and known to be reliable and valid. The tests embrace an appropriate span of cognitive domains and are known to be sensitive to most of the causes of mild cognitive dysfunction. Data establishing the reliability and validity of the CNSVS battery have been published elsewhere (Gualtieri & Johnson, 2006d).

Verbal memory (VBM) and visual memory (VIM) are adaptations of the Rey Auditory Verbal Learning Test and the Rey Visual Design Learning Test (Rey, 1964; Taylor, 1959). VBM and VIM are recognition tests, however, not tests of recall. Correct responses from VBM and VIM are summed to generate a composite memory or memory domain score.

The Finger Tapping Test (FTT) is one of the core tests of the Halstead-Reitan Battery, but similar tests were used by 19th century psychologists like Wundt, Galton, and Cattell. Symbol digit coding (SDC) is based on the Symbol Digit Modalities Test (SDMT) (Smith, 1982), which is a variant of the Wechsler digit symbol substitution test (DSST). The total of right and left taps from the FTT and total correct responses on the SDC generates a composite score for psychomotor speed.

The Shifting Attention Test (SAT) measures the subject’s ability to shift from one instruction set to another quickly and accurately. Other computerized batteries, such as the Neurobehavioral Evaluation System 2 (NES2), CogState, and CANTAB, have SATs. Color-shape tests such as the SAT have been used in cognitive imaging studies (Le, Pardo, & Hu, 1998; Nagahama et al., 1998). A domain score for cognitive flexibility is generated by taking the number of correct responses on the SAT and subtracting the number of errors on the SAT and the ST.

The CPT is a measure of vigilance or sustained attention (Rosvold & Delgado, 1956). A domain score for complex attention is generated by adding the number of errors committed in the CPT, the SAT, and the Stroop. The domain score for vigilance attention is the sum of CPT errors.

A composite score, the Neurocognition Index (NCI), is computed as the average of the z scores of five domains (memory, psychomotor speed, reaction time, complex attention, and cognitive flexibility).

The CNSVS battery has been normed in 1,069 normal volunteers who were in good health without past or present psychiatric or neurological disorders, head injury, learning disabilities, and so on, and free of any centrally acting medications. The subjects ranged in age from 7 to 90. Peak performance on the tests is achieved during the third decade of life and declines gradually thereafter.

### Table 1

Demographic Characteristics of the Five Groups

<table>
<thead>
<tr>
<th>Group</th>
<th>NML</th>
<th>MPH-OROS</th>
<th>ATMX</th>
<th>AMP</th>
<th>ADHD</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>101</td>
<td>51</td>
<td>45</td>
<td>81</td>
<td>95</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>12.35</td>
<td>11.96</td>
<td>13.89</td>
<td>13.33</td>
<td>13.54</td>
</tr>
<tr>
<td>SD</td>
<td>2.82</td>
<td>2.95</td>
<td>2.62</td>
<td>3.28</td>
<td>3.13</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>49</td>
<td>39</td>
<td>27</td>
<td>66</td>
<td>65</td>
</tr>
<tr>
<td>Female</td>
<td>52</td>
<td>12</td>
<td>18</td>
<td>15</td>
<td>30</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>78</td>
<td>33</td>
<td>40</td>
<td>61</td>
<td>83</td>
</tr>
<tr>
<td>Black</td>
<td>10</td>
<td>16</td>
<td>4</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>Asian</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Hispanic</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Native American</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Dose</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>65.2</td>
<td>39.0</td>
<td>23.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>28.6</td>
<td>13</td>
<td>15.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: NML = normal controls; MPH-OROS = Concerta; ATMX = atomoxetine; AMP = Adderall XR.
Test-retest (TRT) reliability of the CNSVS battery was established in a study of 99 Ss, 40 normal volunteers, and 59 psychiatric patients who took the entire battery on two separate occasions, separated on average by 62 days. The TRT interval ranged from 1 to 282 days, with a median interval of 27 days. Reliability coefficients ranged from .65 (Attention) to .87 (Psychomotor Speed). The TRT of the CNSVS battery is comparable to those reported for similar, traditional tests and to similar tests in other computerized test batteries (Gualtieri & Johnson, 2006d).

The concurrent validity of the CNSVS battery was established in a series of studies comparing the performance of subjects on CNSVS to their performance on conventional neuropsychological tests and on another computerized neurocognitive test, the NES2 (Baker et al., 1985). The conventional tests were the Rey Auditory Verbal Learning Test, Logical Memory and Facial Recognition from the Wechsler Memory Test, a mechanical finger tapper, the ST, Trails B, and the Verbal Fluency Test. From the NES2, the comparison tests were Finger Tapping, Switching Attention, and the CPT. CNSVS tests were moderately well correlated with tests of psychomotor speed (finger tapping [.41–.52] and coding [.6–.79]) and executive function on the NES2 (.51–.55). Correlations between the CPT in CNSVS and the NES2 were low (.26–.47). The concurrent validity of the CNSVS battery is comparable to similar conventional neuropsychological tests.

The discriminant validity of the CNSVS battery has been established in studies of patients with mild cognitive impairment (MCI) and early dementia (Gualtieri & Johnson, 2006c); Post-Concussion Syndrome (PCS) and severe traumatic brain injury (Gualtieri & Johnson, 2005, 2006b); ADHD (Gualtieri & Johnson, 2006a); depression (Gualtieri & Johnson, 2006e); schizophrenia and bipolar disorder (Gualtieri & Johnson, 2006a); and malingering (Gualtieri & Johnson, 2006d).

Table 2

<table>
<thead>
<tr>
<th></th>
<th>NML</th>
<th>MPH-OROS</th>
<th>ATMX</th>
<th>AMP</th>
<th>ADHD Controls</th>
<th>MANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>N</td>
<td>101</td>
<td>51</td>
<td>45</td>
<td>81</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>Neurocognition index</td>
<td>99.08</td>
<td>13.29</td>
<td>84.63</td>
<td>23.68</td>
<td>89.46</td>
<td>19.18</td>
</tr>
<tr>
<td>Memory</td>
<td>98.62</td>
<td>16.47</td>
<td>89.71</td>
<td>23.70</td>
<td>90.38</td>
<td>20.89</td>
</tr>
<tr>
<td>Psychomotor speed</td>
<td>100.79</td>
<td>28.28</td>
<td>90.85</td>
<td>28.73</td>
<td>92.11</td>
<td>22.26</td>
</tr>
<tr>
<td>Reaction time</td>
<td>99.89</td>
<td>14.50</td>
<td>96.55</td>
<td>18.25</td>
<td>101.77</td>
<td>17.20</td>
</tr>
<tr>
<td>Cognitive flexibility</td>
<td>99.39</td>
<td>15.32</td>
<td>92.19</td>
<td>24.16</td>
<td>89.60</td>
<td>23.38</td>
</tr>
<tr>
<td>Attention</td>
<td>100.05</td>
<td>16.52</td>
<td>81.00</td>
<td>31.25</td>
<td>84.74</td>
<td>26.75</td>
</tr>
<tr>
<td>Vigilance</td>
<td>100.00</td>
<td>17.47</td>
<td>85.32</td>
<td>29.44</td>
<td>87.08</td>
<td>29.71</td>
</tr>
</tbody>
</table>

Note: NML = normal controls; MPH-OROS = Concerta; ATMX = atomoxetine; AMP = Adderall XR; MANOVA = multivariate analysis of the variance.

Method

The CNSVS database contains records from more than 4,000 patients with neurological and/or psychiatric disorders. The database was scanned for patients who met the following criteria: (a) primary diagnosis ADHD; (b) age, 10 to 18 years; (c) no comorbid neurological conditions, cognitive disorders (e.g., learning disabilities, brain injury), or psychiatric disorders (e.g., anxiety, depression, autism); (d) treatment with AMP, ATMX, or MPH-OROS, stable doses maintained for at least 4 weeks; and (e) no concurrent medications.

This process identified four groups: Three groups were children and adolescents treated with MPH-OROS, ATMX, or AMP. The charts were reexamined by the authors to establish that parents and treating clinicians agreed that the patient was a positive responder to the medication and that no further treatment recommendations were contemplated. The fourth group was untreated and newly diagnosed ADHD patients. These were the untreated ADHD controls.

The NML were selected from the CNSVS normative database. These were individuals with no present or past psychiatric, developmental, or neurological disorder, in good health, in good standing in school, and taking no current medications. All of the patients had been tested within a 14-month period (July 2003–August 2004).

Results

Subject data, test scores, and results of the multivariate analysis are presented in Table 2. There were no significant
differences among the groups in terms of race, but there were for age and gender; therefore, the multivariate analysis of the variance (MANOVA) incorporate age and gender as covariates.

The results in Table 2 indicate highly significant group differences among the five groups, in all six cognitive domains and in the summary score, NCI. The next step was to determine the source of the differences.

It did not reside in the between-drug comparisons (Pillai’s trace, $F = 1.053$, $p > .40$). In light of that, the results of the three drug groups were combined, and analysis was performed comparing three groups, NML, medicated ADHD patients, and unmedicated ADHD controls. As expected, group differences, controlling for age and gender, were significantly different for the NCI and all six domains (Table 3).

The data in Table 4 indicate where the significant group differences reside. Analysis of variance (ANOVA) with Bonferroni correction measured the significance of differences, and effect sizes were measured by Cohen’s $d$. The most impressive differences were, as expected, between NML and untreated ADHD; significant differences and moderate-to-strong effect sizes were observed in every domain, save memory, and in the NCI. The NML were also significantly superior to treated ADHD patients in every domain, except reaction time and psychomotor speed. Medicated patients differed from untreated patients in the domains of reaction time, cognitive flexibility, and vigilance attention, although the effects were small.

The data are presented graphically in Figure 1.

**Discussion**

ADHD treatment, conducted by experienced practitioners in a clinic specializing in ADHD and other neuropsychiatric disorders, does not necessarily normalize the cognitive performance of ADHD children and adolescents. Taking the NCI as a benchmark, untreated ADHD patients perform 15% lower than normals. However, treated ADHD patients perform 10% lower than normals.

Three alternative conclusions might be drawn from these data. First, it is possible that the treatment was suboptimal. Perhaps better practitioners might have achieved better results. Second, because ADHD is a behavioral disorder, cognitive performance is irrelevant to treatment. We do not accept either of these. We believe that the third possible conclusion is the right one: That even with optimal treatment by experienced clinicians, children, and adolescents with ADHD, as a group, have persistent cognitive disabilities. Drug treatment may ameliorate the cognitive difficulties of ADHD patients, but it does not necessarily normalize cognition. Medication treatment may be necessary, therefore, for children and adolescents with ADHD, but it may not be sufficient.

This conclusion would be more convincing if we had systematic behavioral data to present. The premise on which the study is based, after all, is that the treated subjects were positive responders on clinical grounds. If we had objective data, from behavioral observations, or even subjective data from rating scales, which would lend strength to the assertion that all the subjects were in fact responders. We regret that no such data were available. On the other hand, in the neuropsychiatry clinics, medication response in ADHD children and adolescents is assessed on the basis of interviews with parents and patients, reports from teachers and improvement in grades, and by comparing different drugs and different doses. This is, in fact, the way response is measured in clinical practice, and the point of this study is simply this: Under the usual circumstances of good clinical care,
medication treatment does not necessarily resolve all of the cognitive problems that ADHD patients have.

The patients in this study were treated with medication because they were doing poorly in school. With treatment, their performance in school improved. We regret that we do not have grades to analyze and report and to correlate with improvement in cognition on the computerized neurocognitive test battery. That would be the appropriate subject for a prospective study, and would address the ecological validity of the cognitive testing. The degree to which neuropsychological tests correlate with real-world performance has been the subject of a good deal of research and controversy. Overall, the research suggests that many neuropsychological tests have a moderate level of ecological validity when predicting everyday cognitive functioning. The strongest relationships are noted when the outcome measure corresponds to the cognitive domain assessed by the neuropsychological tests (Chaytor & Schmitter-Edgecombe, 2003).

### Table 4

<table>
<thead>
<tr>
<th></th>
<th>Normals Versus Treated</th>
<th></th>
<th>Normals Versus ADHD Controls</th>
<th></th>
<th>Treated Versus ADHD Controls</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ANOVA Bonferroni</td>
<td>Cohen’s d</td>
<td>ANOVA Bonferroni</td>
<td>Cohen’s d</td>
<td>ANOVA Bonferroni</td>
<td>Cohen’s d</td>
</tr>
<tr>
<td>Neurocognition index</td>
<td>0.0001</td>
<td>0.559</td>
<td>0.0000</td>
<td>0.888</td>
<td>0.2435</td>
<td>0.209</td>
</tr>
<tr>
<td>Memory</td>
<td>0.0072</td>
<td>0.423</td>
<td>0.1214</td>
<td>0.335</td>
<td>1.0000</td>
<td>–0.089</td>
</tr>
<tr>
<td>Psychomotor speed</td>
<td>0.0848</td>
<td>0.271</td>
<td>0.0012</td>
<td>0.551</td>
<td>0.1819</td>
<td>0.261</td>
</tr>
<tr>
<td>Reaction time</td>
<td>1.0000</td>
<td>0.058</td>
<td>0.0009</td>
<td>0.551</td>
<td>0.0008</td>
<td>0.442</td>
</tr>
<tr>
<td>Cognitive flexibility</td>
<td>0.0422</td>
<td>0.336</td>
<td>0.0001</td>
<td>0.773</td>
<td>0.0386</td>
<td>0.309</td>
</tr>
<tr>
<td>Attention</td>
<td>0.0002</td>
<td>0.585</td>
<td>0.0000</td>
<td>0.936</td>
<td>0.2476</td>
<td>0.22</td>
</tr>
<tr>
<td>Vigilance</td>
<td>0.0242</td>
<td>0.388</td>
<td>0.0000</td>
<td>0.878</td>
<td>0.0014</td>
<td>0.431</td>
</tr>
</tbody>
</table>

Note: ANOVA = analysis of variance.

### Figure 1

Normals, Medicated Patients, and Untreated Patients

Note: NML = normal controls; MED = ADHD patients on medication; ADHD = unmedicated ADHD patients; NCI = neurocognition index; MEM = memory domain score; PMS = psychomotor speed; RT = reaction time; CF = cognitive flexibility; ATT = complex attention; AV = vigilance attention.
The relevance of neuropsychological tests to ADHD diagnosis and to treatment assessment has also been the subject of some concern. The results of psychological tests, for example, the freedom from distractibility factor of the Wechsler Intelligence Scale for Children (WISC-III) is not found to be “a reliable or a valid index of attention or a diagnostic screening measure for identifying children with ADHD” (Reinecke, Beebe, & Stein, 1999, p. 322). The correlation between performance on the CPT and parent or teacher rating scales is modest at best (Forbes, 1998; McGee, Clark, & Symons, 2000; Raggio & Pierce, 1999; Rielly, Cunningham, Richards, Elbard, & Mahoney, 1999), and computerized CPTs like the tests of variables of attention (TOVA) generate unacceptably high false positive rates (30%) in NML and children with other psychiatric disorders (28%) (Forbes, 1998; Schatz, Ballantyne, & Trauner, 2001).

If a child’s behavior, then, is improved when he or she has been treated with an ADHD medication, and if grades improve as well, then why should it matter if performance on a battery of neuropsychological tests is 10% lower than normal? We believe that this is more than a trivial result and that it does have clinical relevance. Even with successful treatment, ADHD children and adolescents have subtle but clearly demonstrable cognitive deficits. Some of these patients may need more than just medication. Furthermore, it is possible that the insufficiency of ADHD treatment is responsible for the relatively low adherence rates over time. Recent reports indicate that one third of ADHD children are no longer taking medication after 2 years (Bussing et al., 2005) and that only about one in five take medication over a longer period (Miller, Lalonde, & McGrail, 2004).

Appendix

The CNS Vital Signs (CNSVS) Battery

Verbal Memory (VBM) Test and Visual Memory (VIM) Test

Vital Signs includes parallel tests of verbal memory (word list learning) and visual memory (figure learning). The tests are virtually identical, but one uses words as stimuli, the other geometric shapes.

The VBM test is an adaptation of the Rey Auditory Verbal Learning Test (Rey, 1964; Taylor, 1959). It is a recognition test, however, not a test of recall. In the CNSVS version, 15 words are presented, one by one, on the screen. A new word is presented every 2 s. The subject is asked to remember these words. Then a list of 30 words is presented. The 15 target words are mixed randomly among 15 new words. When the subject recognizes a word from the original list, he or she presses the space bar. After this trial of 30 stimuli, the subject goes on to do the next six tests. At the end of the battery, about 20 min later, the 15 target words appear again, mixed with 15 new nontarget words.

The VIM Test is based on the Rey Visual Design Learning Test; the latter is in turn, a parallel to the Rey Auditory Verbal Learning Test, using geometric figures rather than words, and requiring the subject to draw the figures from memory. In CNSVS, the VIM Test is just like the VBM Test. In the VIM Test, 15 geometric figures are presented; the subject has to identify those figures nested among 15 new figures. Then, after five more tests, there is a delayed recognition trial.

The VBM draws from a reservoir of 120 words selected from word-frequency tables. The VIM draws from a reservoir of 120 simple geometric designs. The scoring is correct hits and correct passes, immediate and delayed. Correct responses from VBM and VIM are summed to generate a composite memory or memory domain score. The highest score one can attain is 120, the lowest is 60. Scores below 60 suggest willful exaggeration.

Finger Tapping Test (FTT)

The FTT is one of the most commonly used tests in neuropsychology, because of its simplicity and reliability, and also generates relevant data about fine motor control, which is based on motor speed as well as kinesthetic and visual-motor ability (Mitrushina, Boone, & D’Elia, 1999). It was one of the core tests of the Halstead-Reitan Battery, which dates to the 1940s, but similar tests were used by 19th-century psychologists like Wundt, Galton, and Cattell. The FTT is believed to be one of the most sensitive neuropsychological tests for determining brain impairment (Mitrushina et al., 1999).

In CNSVS, the FTT is a very simple test. Subjects are asked to press the space bar with their right index finger as many times as they can in 10 s. They do this once for practice, and then there are three test trials. The test is repeated with the left hand. The score is the average number of taps, right and left.

Symbol digit coding (SDC)

The Symbol Digit Modalities Test (SDMT) (Smith, 1982) is a variant of the Wechsler DSST, but the position of symbols and digits is reversed. The clinical and psychometric properties of the SDMT are similar to those of the DSST. Although the SDMT may be a harder test, and
thus more sensitive to neurotoxicity, performance on the SDMT and the DSST are highly correlated (Lezak, 1994). Smith maintained that the SDMT was “usually the most sensitive (test) to the presence of acute or chronic ‘organic’ cerebral dysfunction” (Smith, 1982).

In the CNSVS SDC, the subject is given a training session to learn how to link numbers to digits. The test itself consists of serial presentations of screens, each of which contains a bank of eight symbols above and eight empty boxes below. The subject types in the number that corresponds to the symbol that is highlighted. Only the digits from 2 through 9 are used; this is to avoid confusion between “1” and “I” on the keyboard. The test lasts for 120 s. The goal is to type in as many correct numbers as one can in 120 s.

Neither the SDMT nor the DSST are suitable for repeated administration, because subjects are able to remember the code and thus accelerate their performance (Hindmarch, 1980). Modifications in the test are necessary if it is to be used repeatedly; for example, changing the code in a random way on successive administrations. The SDC in CNSVS draws from a reservoir of 32 symbols. Each time the test is administered, the program randomly chooses eight new symbols to match to the eight digits.

Scoring is the number of correct responses generated in 2 min. The total of right and left taps from the FTT and total correct responses on the SDC generates a composite score for psychomotor speed.

The Stroop Test (ST)

There have been several versions of the ST over the years. The modification adopted for CNSVS uses only four colors/color words (red, green, yellow, blue), and only one key is in play, the space bar. The test has three parts. In the first, the words RED, YELLOW, BLUE, and GREEN (printed in black) appear at random on the screen, and the subject presses the space bar as soon as he or she sees the word. This generates a simple reaction time score.

In the second part, the words RED, YELLOW, BLUE, and GREEN appear on the screen, printed in color. The subject is asked to press the space bar when the color of the word matches what the word says. This generates a complex reaction time score.

In the third part, the words RED, YELLOW, BLUE, and GREEN appear on the screen, printed in color. The subject is asked to press the space bar when the color of the word does not match what the word says. This part also generates a complex reaction time score, called the color-word reaction time. The color-word reaction time is on average 120 ms longer than the complex reaction time generated in part two of the test (range, 78–188 ms) (the Stroop effect). Part three also generates an error score.

Averaging the two complex reaction time scores from the ST generates a domain score for reaction time. It might be more precise to refer to this domain as information processing speed.

The Shifting Attention Test (SAT)

The SAT measures the subject’s ability to shift from one instruction set to another quickly and accurately. In the SAT, subjects are instructed to match geometric objects either by shape or by color. Three figures appear on the screen, one on top and two on the bottom. The top figure is either a square or a circle. The bottom figures are a square and a circle. The figures are either red or blue; the colors are mixed randomly. The subject is asked to match one of the bottom figures to the top figure. The rules change at random. For one presentation, the rule is to match the figures by shape, for another, by color. This goes on for 90 s. The goal is to make as many correct matches as one can in the time allotted. The scores generated by the SAT are correct matches, errors, and response time (ms). A domain score for cognitive flexibility is generated by taking the number of correct responses on the SAT and subtracting the number of errors on the SAT and the ST.

There is not a precise parallel to the SAT in the compendium of conventional neuropsychological tests, although Trails B and the Wisconsin Card Sort are sometimes considered to be tests of shifting attention. Computerized tests, however, such as the NES2, CogState, and CANTAB, have SATs that are not dissimilar to the SAT, and color-shape tests like the SAT have been used in cognitive imaging studies (Le et al., 1998; Nagahama et al., 1998).

The Continuous Performance Test (CPT)

The CPT is a measure of vigilance or sustained attention or attention over time (Rosvold & Delgado, 1956). It has been a popular test because of its robust relationship to psychiatric disorders. Poor performance on the CPT has been reported in ADHD (Epstein, Johnson, Varia, & Conners, 2001; Sykes, Douglas, Weiss, & Minde, 1971), learning disabilities (Lindsay, Tomazic, Levine, & Accardo, 2001; McGee et al., 2000), patients with epilepsy (Mirkys & van Buren, 1965), and schizophrenia (Vadhan, Serper, Harvey, Chou, & Cancro, 2001; Wohlberg & Kornetsky, 1973). It is sensitive to CNS dysfunction in general, and is not specific to any particular condition (Riccio & Reynolds, 2001).
The CPT is also sensitive, for better or worse, to the effects of various drugs. In ADHD children, performance on the CPT is reliably improved by stimulant medications (Barkley, 1977; Riccio, Waldrop, Reynolds, & Lowe, 2001). Alcohol consumption (Dougherty, Marsh, Moeller, Chokshi, & Rosen, 2000) adversely affects performance on the CPT, but nicotine tends to improve performance on the test (Levin, Conners, Silva, Canu, & March, 2001). Certain anticonvulsant medications impair performance on the CPT (Hutt, Jackson, Belsham, & Higgins, 1968).

The CPT in Vital Signs is a conventional version of the test, although it is shorter than some other versions. In the Vital Signs CPT, the subject is asked to respond to target stimulus “B” but not to any other letter. In 5 min, the test presents 200 letters; 40 of the stimuli are targets (the letter “B”) and 160 are nontargets (other letters). The stimuli are presented at random, although the target stimulus is blocked so it appears 8 times during each minute of the test.

Scoring is correct responses, commission errors (impulsive responding), and omission errors (inattention). The CPT also reports subjects’ choice reaction time for each variable. A domain score for complex attention is generated by adding the number of errors committed in the CPT, the SAT, and the Stroop.

References


C. Thomas Gualtieri, MD, is medical director at the North Carolina Neuropsychiatry Clinics, Chapel Hill and Charlotte.

Lynda G. Johnson, PhD, is director of neuropsychology at the North Carolina Neuropsychiatry Clinics. Drs Gualtieri and Johnson are two of the developers of the CNS Vital Signs computerized test battery.