Identifying Cognitive Problems in Children and Adolescents with Depression Using Computerized Neuropsychological Testing

Brian L. Brooks
Neurosciences, Alberta Children’s Hospital, Calgary, Canada and Departments of Pediatrics and Clinical Neurosciences, University of Calgary, Calgary, Canada

Grant L. Iverson
Department of Psychiatry, University of British Columbia, Vancouver, Canada and British Columbia Mental Health & Addiction Services, Coquitlam, Canada

Elisabeth M. S. Sherman
Neurosciences, Alberta Children’s Hospital, Calgary, Canada and Departments of Pediatrics and Clinical Neurosciences, University of Calgary, Calgary, Canada

Marie-Claude Roberge
British Columbia Mental Health & Addiction Services, Coquitlam, Canada and University of Ottawa, Ottawa, Canada

Depression in children and adolescents can negatively impact cognitive functioning, social development, and academic performance. The purpose of this study was to determine whether a computerized battery of neuropsychological tests could detect neurocognitive difficulties in children and adolescents with depression. Participants included 30 children and adolescents between the ages of 9 and 17 years ($M = 14.6$, $SD = 2.1$) with a clinical diagnosis of depression. Healthy control participants were individually matched on age, education, sex, race, primary language, handedness, and self-reported computer familiarity. All participants completed the Central Nervous System Vital Signs computerized battery. This battery of seven tests yields 23 test scores and 5 domain scores (Memory, Psychomotor Speed, Reaction Time, Complex Attention, and Cognitive Flexibility). Children and adolescents with depression performed worse on the Memory (Cohen’s $d = .43$) and Complex Attention domains ($d = .58$) than matched controls. On the individual test scores, children and adolescents with depression performed worse on delayed verbal memory ($d = .63$), delayed visual memory ($d = .34$), measures of reaction time ($d = .34–.53$), and accuracy/inhibition on complex attention tasks ($d = .49–.65$). When considering the five domain scores simultaneously, children and adolescents with depression were more likely to have two or more scores at or below the 5th percentile ($p = .05$). Children and adolescents with depression have problems with reduced processing speed, memory for verbal
information, and executive functioning on this computerized battery of tests, which represents a feasible method for neuropsychological screening.

Key words: adolescent, children, cognition, computerized testing, depression, memory

INTRODUCTION

Depression is associated with perceived and/or measurable problems in cognition. Clinicians are interested in identifying the presence of cognitive problems as part of the evaluation, diagnosis, and treatment of depression. The *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition, Text Revision (American Psychiatric Association, 2000) indicates that, “many individuals report impaired ability to think, concentrate, or make decisions (Criterion A8). They may appear easily distracted or complain of memory difficulties. Those in intellectually demanding academic or occupational pursuits are often unable to function adequately even when they have mild concentration problems…” (p. 350).

Cognitive problems associated with depression have been well studied in adults, with meta-analyses identifying reductions in sustained attention, psychomotor speed, memory, verbal fluency, and cognitive flexibility as the most prominent cognitive features (Burt, Zembar, & Niederehe, 1995; Christensen, Griffiths, Mackinnon, & Jacomb, 1997; Henry & Crawford, 2005; Zakzanis, Leach, & Kaplan, 1998). Less research has been conducted with children and adolescents with depression. The neuropsychological effects of depression in children and adolescents have included measurable problems across most domains, including Sustained Attention, Processing Speed, Reaction Time, Learning and Memory, Sequencing, and Problem-Solving (Cataldo, Nobile, Lorusso, Battaglia, & Molteni, 2005; Emerson, Mollet, & Harrison, 2005; Günther, Holtkamp, Jolles, Herpertz-Dahlmann, & Konrad, 2004; Kovacs & Goldston, 1991; Lauer et al., 1994; Wilkinson & Goodyer, 2006). Clearly, identifying the presence of cognitive problems in children and adolescents with depression is important because impaired thinking can have a negative impact on academic abilities (Fröjd et al., 2008; Kovacs & Goldston, 2004) and social functioning (Kovacs & Goldston, 2004).

The assessment of cognition in children and adolescents with depression using computerized neurocognitive batteries likely represents a feasible option for screening cognitive abilities. For example, Kyte, Goodyer, and Sahakian (2005) reported that adolescents with first-episode depression had worse performance on several measures of executive functioning on the Cambridge Neuropsychological Test Automated Battery.

The Central Nervous System (CNS) Vital Signs computerized neurocognitive battery (Gualtieri & Johnson, 2006d) is normed across the lifespan (i.e., ages 7 to 90) and represents one computerized method for evaluating cognition. Recent studies have demonstrated that this computerized test battery can rapidly identify cognitive problems in adults with various psychiatric diagnoses, including depression (Gualtieri & Johnson, 2006a, 2007, 2008; Gualtieri, Johnson, & Benedict, 2006; Gualtieri & Morgan, 2008; Iverson, Brooks, Young, Johnson, & Gualtieri, 2007c), anxiety (Gualtieri & Morgan), bipolar disorder (Gualtieri & Johnson, 2006b; Gualtieri & Morgan; Iverson, Brooks, & Young, in press), and attention deficit/hyperactivity disorder (ADHD; Iverson, Brooks, Weiss, Johnson, & Gualtieri, 2007b). The use of this computerized battery in children and adolescents with various clinical diagnoses is less well known, although there have been a few studies examining performance in children and adolescents with treated (Gualtieri & Johnson, 2006c, 2008) and untreated (Iverson, Brooks, Weiss, Gualtieri, & Johnson, 2007a) ADHD.

The purpose of this study is to examine whether the CNS Vital Signs computerized test battery is sensitive to cognitive problems associated with depression in children and adolescents. This is a preliminary examination of the sensitivity of CNS Vital Signs in this clinical population because of the relatively small sample size and the absence of clearly defined depression variables. It is hypothesized that children and adolescents with depression will have worse cognitive abilities, particularly attention, processing speed, memory, and flexible thinking, than matched controls (i.e., these are the cognitive domains measured by the CNS Vital Signs battery). In addition, previous work with mood disorders and performance on the computerized battery (Iverson et al., 2007c; Iverson, Brooks, & Young, in press) would suggest there is a subset of patients with depression who have substantial cognitive problems, and this subset accounts for the group differences and effect sizes.

METHOD

Participants

The participants for this study included 30 children and adolescents, selected from an archival database, who were diagnosed with depression (i.e., major depressive disorders).
disorder or depressive disorder not otherwise specified). Clinicians at the North Carolina Neuropsychiatry Clinics (Chapel Hill and Charlotte, North Carolina) gave a primary diagnosis of depression to all patients according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (American Psychiatric Association, 1994). The majority of the sample were taking antidepressant medication at the time of their evaluation (e.g., 63.3%), and 16.7% had a comorbid anxiety disorder.

The average age of the patients was 14.6 years ($SD = 2.1$, range = 9–17, median = 15), and their average years of education was 9.2 ($SD = 2.2$, range = 4–12, median = 9). The majority of the patient sample were right-handed ($n = 29$, 96.7%), female ($n = 22$, 73.3%), and Caucasian ($n = 28$, 93.3%; African American, $n = 2$, 6.7%). The entire sample identified English as their primary language and reported they were still going to school. The sample either had some ($n = 4$; 13.3%) or frequent ($n = 26$; 86.7%) computer use.

Children and adolescents with depression were compared to 30 healthy control children and adolescents selected from the CNS Vital Signs normative database. The CNS Vital Signs database contains 1,069 healthy people between the ages of 7 and 90 years who comprised the standardization sample. They were screened to be in good health, without past or present psychiatric or neurological disorders (i.e., head injury, learning disabilities, etc.), and were free of any centrally acting medications. The control participants were individually and closely matched to the participants with depression on education ($M = 9.3$, $SD = 2.1$; $t(58) = 12$, $p = .90$), and individually and precisely matched on age ($t(58) = .90$, $p = 1.00$), sex, race, primary language, handedness, student status, and self-reported computer use. The precise matching procedure controls for variability in neurocognitive test performance that might result from differences in demographic variables (for a discussion of the impact of computer familiarity on test performance, see Iverson, Brooks, Ashton, Johnson, & Gualtieri, 2009).

Measures

CNS Vital Signs is a computerized battery composed of seven common neuropsychological measures, including verbal and visual memory, finger tapping, Symbol Digit Coding, the Stroop test, a Shifting Attention Test, and a Continuous Performance Test (CPT). The battery generates 15 primary scores (as well as several secondary scores), which are used to calculate 5 domain scores (Memory, Psychomotor Speed, Reaction Time, Cognitive Flexibility, and Complex Attention). The domain (or index) scores are converted into standard scores and are presented with a mean of 100 and a standard deviation (SD) of 15.

The battery is normed across the lifespan for children, adolescents, and adults, and it is presented at a grade-four reading level. The measures have adequate test-retest reliability (Pearson’s $r$ for the domain scores ranged from $r = .65–.87$; mean interval of 62 days, range = 1–282 days), adequate concurrent validity with traditional paper-and-pencil measures and other computerized tests (i.e., most of the Pearson’s $r$ for the selected scores were small to medium for both the traditional paper-and-pencil and other computerized tests), and the domain scores have been shown to discriminate between various clinical groups. More detailed information about the reliability and validity of this test battery is available from other sources (Gualtieri & Johnson, 2005, 2006b; Gualtieri et al., 2006). Descriptions of the seven measures on the CNS Vital Signs battery are presented in Table 1.

Analyses

Analysis of the CNS Vital Signs test results involved: a) comparing the two groups on the five domain scores using multivariate analyses of variance (MANOVA) and following up with independent samples $t$ tests; b) examining the effect sizes for the individual primary scores; and c) comparing the base rates of low domain scores using nonparametric analyses (i.e., chi-square; $X^2$). Calculations for the base rates of low scores involve simultaneously examining performance on the five domain scores, rather than performance on each domain in isolation. The base rates of low domain scores were calculated using four cut-off scores that might be routinely used in clinical practice, including: a) below the 16th percentile (i.e., more than 1 SD below the mean or index $< .85$), b) below the 10th percentile (i.e., index $< .81$), c) at or below the 5th percentile (i.e., index $\leq .76$), and d) at or below the 2nd percentile (i.e., more than 2 SDs below the mean or index $\leq .70$).

RESULTS

The two groups were compared on the five domain scores using MANOVA followed by $t$ tests. Box’s M test was significant, indicating that the covariance matrices differed ($p = .006$). Moreover, Levene’s test was significant for one of the five index scores (i.e., Complex Attention), indicating heterogeneity of variance between groups. However, MANOVA is generally robust to violations of the general linear model assumptions and will therefore be interpreted. The multivariate effect was significant (Wilks’ Lambda = .82; $F(5, 54) = 2.41$, $p < .05$, partial eta squared $= .18$, observed power = .72).

The follow-up $t$ tests revealed significantly worse neuropsychological test scores for those in the depression
TABLE 1

Descriptions of the CNS Vital Signs Measures

<table>
<thead>
<tr>
<th>CNS Vital Signs Measure</th>
<th>Description of Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal Memory</td>
<td>This test measures recognition memory for words. Fifteen words are presented, 1 by 1, on the screen every 2 seconds. For immediate recognition, the participant has to identify those figures nested among 15 new figures. Then after six more tests, there is a delayed recognition trial.</td>
</tr>
<tr>
<td>Visual Memory</td>
<td>This test measures recognition memory for figures. Fifteen geometric figures are presented, 1 by 1, on the screen. For immediate recognition, the participant has to identify those figures nested among 15 new figures. Then after five more tests, there is a delayed recognition trial.</td>
</tr>
<tr>
<td>Finger Tapping</td>
<td>Participants are asked to press the space bar with their right index finger as many times as they can in 10 seconds. They do this once for practice, and then there are three test trials. The test is repeated with the left hand.</td>
</tr>
<tr>
<td>Symbol Digit Coding</td>
<td>The test consists of serial presentations of screens, each of which contains a bank of eight symbols above and eight empty boxes below. The participant types in the number that corresponds to the symbol that is highlighted. Only the digits from two through nine are used; this is to avoid the confusion between “1” and “I” on the keyboard. Moreover, the participant is only allowed to use the numbers two through nine at the top of a traditional keyboard (i.e., the computer program does not allow a person to use a numerical pad). This prevents the potential for a distinct advantage for those who are skilled at using the numerical pad or for those who are right-versus left-handed.</td>
</tr>
<tr>
<td>Stroop Test</td>
<td>The test has three parts. In the first part, the words RED, YELLOW, BLUE, and GREEN (printed in black) appear at random on the screen, and the participant presses the space bar as soon as he or she sees the word. In the second part, the words RED, YELLOW, BLUE, and GREEN appear on the screen, printed in color. The participant is asked to press the space bar when the color of the word matches what the word says. In the third part, the words RED, YELLOW, BLUE, and GREEN appear on the screen, printed in color. The participant is asked to press the space bar when the color of the word does not match what the word says.</td>
</tr>
<tr>
<td>Shifting Attention Test</td>
<td>A measure of ability to shift from one instruction set to another quickly and accurately. Participants 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 (mixed randomly). The participant is asked to match one of the bottom figures to the top figure. The rules change at random (i.e., match the figures by shape; for another, by color).</td>
</tr>
<tr>
<td>Continuous Performance</td>
<td>A measure of vigilance or sustained attention or attention over time. The participant is asked to respond to the target stimulus “B” but not to any other letter. In 5 minutes, the test presents 200 letters. Forty of the stimuli are targets (the letter “B”), 160 are non-targets (other letters). The stimuli are presented at random, although the target stimulus is “blocked,” so it appears eight times during each minute of the test.</td>
</tr>
</tbody>
</table>

This study sought to determine whether the CNS Vital Signs computerized battery, which is a rapid screen of a person’s thinking abilities, is sensitive to cognitive impairment in children and adolescents with depression. In the present study, children and adolescents with...
depression had statistically lower performance compared with matched controls on the Complex Attention domain (medium effect size, $d = .58$). The Complex Attention domain is a composite measure of a person’s accuracy (i.e., number of errors) during tests of sustained attention, shifting attention, and inhibition. In other words, this domain represents a person’s performance on computerized measures of executive functions. The children and adolescents with depression had worse performance than controls and medium effect sizes on all of the measures that contribute to the Complex Attention domain score: Stroop errors ($d = .56$), Shifting Attention Test errors ($d = .59$), Continuous Performance Test omissions ($d = .57$), Continuous Performance Test commissions ($d = .65$). They also had worse performance and a medium effect size on the Stroop Simple Reaction Time score ($d = .53$). There were no statistically significant differences on the other domain scores, although there was a small-to-medium effect size on the Memory domain ($d = .43$), which was likely the result of lower performance on the delayed Verbal Memory test (medium effect size, $d = .63$) and the delayed Visual Memory test (small effect size, $d = .34$). The absence of statistically significant findings, despite having small-to-medium effect sizes, is likely the result of the relatively small samples for the two groups and the increased heterogeneity of test scores in the children with depression.

Overall, the results of this study suggest that children and adolescents with depression have problems with reduced processing speed, memory for verbal information, and set shifting, impulsivity, and inhibition during tests of executive functioning. These results are consistent with existing studies involving traditional paper-and-pencil testing of children and adolescents with depression (e.g., Cataldo et al., 2005; Emerson et al., 2005; Güntner et al., 2004; Kovacs & Goldston, 1991; Lauer et al., 1994; Wilkinson & Goodyer, 2006) and a study involving computerized testing of children and
adolescents with depression (e.g., Kyte et al., 2005) that identified measurable problems with sustained attention, processing speed, reaction time, learning and memory, sequencing, inhibition, and problem-solving.

This study also presented psychometrically derived information on the base rates of low scores, which can be used with this computerized testing to supplement clinical judgment, for identifying those children and adolescents with depression who have cognitive problems. When considering the prevalence of low domain scores, our previous research has often indicated that having two or more domain scores at or below the 5th percentile is suggestive of frank cognitive impairment (e.g., Iverson et al., 2007c; Iverson, Brooks, & Young, in press). In the present study, 30% of children and adolescents with depression had two or more domain scores at or below the 5th percentile. This study supports the hypothesis that there is a subset of patients with depression who have frank cognitive impairment, and it is this subset that drives the effect sizes in group comparisons. Future studies should focus on the differences (e.g., number and severity of depressive symptoms, daily functioning, academic performance, and quality of life) between those with and those without impairment in their thinking abilities.

There were some limitations with the present study that are worth noting. First, the sample sizes of the groups in this study were relatively small, and this likely resulted in low power for statistical analyses. This is most obvious with some of the between-group comparisons that were non-significant but had medium effect sizes. For example, the difference between the children and adolescents with depression and the healthy controls on the Symbol Digit Errors score was not significant ($p = .07$), despite having a medium effect size ($d = .49$). Another limitation of this study is the lack of clinical information on the sample of children and adolescents with depression. For example, we are unable to determine the severity and/or the duration of depression in these patients. We are unable to determine how long a subset of the patients with depression has been on medication, the name of the medication, and whether they have responded to the medication. Moreover, we are unable to determine if there is a relation between more symptoms of depression and worse cognition. The absence of some of this clinical information, which is the result of the data being obtained from an archival database, does not discount the new knowledge obtained from this study. It does, however, limit the generalizability of the findings, and further evaluations with larger, better-defined clinical groups are warranted.

A third limitation of this study is that we do not have information on the subjective report of cognitive problems, and we do not know how many children and adolescents also have problems with social and academic skills. This will be an important contribution of future research using brief computerized screening batteries of cognition in children and adolescents with depression.

The CNS Vital Signs computerized battery represents a feasible methodology for screening patients’ cognitive abilities. By quickly getting a sense of whether a patient with depression has cognitive impairment, a clinician can consider the most appropriate treatment and determine if a more thorough neuropsychological assessment is warranted.

**ACKNOWLEDGEMENTS**

The authors thank Drs. Gualtieri and Johnson for providing the data for this study. This study was presented at the 28th annual conference of the National Academy of Neuropsychology (October 2008).

**REFERENCES**


