Validation of the adapted CogState brief battery in hematopoietic cell transplant patients

Bethany Gourley
Validation of the Adapted CogState Brief Battery in Hematopoietic Cell Transplant Patients

by

Bethany Gourley

Thesis

Submitted to the Department of Psychology

Eastern Michigan University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Clinical Psychology

Thesis Committee:

Flora Hoodin, PhD, Chair

Renee Lajiness-O’Neill, PhD

Karen Saules, PhD

July 25, 2013

Ypsilanti, Michigan
Abstract

A small literature has documented cognitive deficits in adult hematopoietic cell transplant (HCT) survivors across the transplant trajectory, primarily occurring in memory, executive function, attention, and processing speed. Although HCT-associated cognitive decline occurring within one year of transplantation is well documented, only two studies have longitudinally investigated cognitive function in HCT survivors beyond one year. Furthermore, studies demonstrating neuropsychological decline have made use of numerous measures and varying impairment criteria, making the compilation of findings across studies challenging. Another difficulty with the current literature base is the use of traditional neuropsychological tests that are susceptible to practice effects and thus may be inappropriate for repeated assessment at short intervals. The present study aimed to evaluate the construct and criterion validity of the Adapted CogState Brief Battery, a computerized measure that has demonstrated construct and criterion validity in medical populations with minimal practice effects and takes nominal time to administer.

Participants were HCT survivors (n = 20) and their spousal partners (n = 20). They completed two brief neuropsychological batteries: CogState and traditional, counterbalanced to obviate order effects. Overall, relationships between CogState and traditional neuropsychological tasks were weaker than predicted, although some task relationships trended toward significance in predicted directions. Further, the battery demonstrated poor criterion validity in that most tasks did not discern HCT survivors from healthy spousal partners, with the exception of Identify, which identified worse performance in HCT survivors. In general, results showed promise for the use of various CogState tasks as screening tools for evaluating cognitive change over time. Future research evaluating the efficacy of CogState and traditional measures of neuropsychological function in HCT survivors is warranted.
# Table of Contents

Abstract ........................................................................................................................................ ii

List of Tables .................................................................................................................................. viii

List of Figures ................................................................................................................................... ix

Introduction ........................................................................................................................................ 1

HCT Overview ................................................................................................................................. 2

Pre-HCT Procedures ....................................................................................................................... 3

HCT Transplantation ....................................................................................................................... 3

Post-HCT Physical complications .................................................................................................... 4

Post-HCT Psychological complications ........................................................................................... 5

Neuropsychological Risks ............................................................................................................... 6

Cross-Sectional Studies .................................................................................................................. 11

Pre-HCT ........................................................................................................................................... 11

Six months post-HCT ..................................................................................................................... 14

One to four years post-HCT ........................................................................................................... 15

Prospective Studies ........................................................................................................................ 18

Pre-HCT through 3 months post-HCT ........................................................................................... 18

Pre-HCT through 8 months post-HCT ........................................................................................... 19

Pre-HCT through 1 year post-HCT ................................................................................................. 20

Pre-HCT through 18 months post-HCT ........................................................................................ 26

Pre-HCT through 5 years post-HCT ............................................................................................... 27

Delerium in HCT patients ............................................................................................................... 29

Other Factors Associated with Cognitive Decline .......................................................................... 33
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risks associated with HCT pre-conditioning regimens</td>
<td>33</td>
</tr>
<tr>
<td>GVHD-related cognitive deficits</td>
<td>34</td>
</tr>
<tr>
<td>Neuropsychological Domains Affected</td>
<td>35</td>
</tr>
<tr>
<td>Alternatives to Traditional Neuropsychological Measurement</td>
<td>36</td>
</tr>
<tr>
<td>Subjective cognitive self-evaluation</td>
<td>37</td>
</tr>
<tr>
<td>Computerized cognitive testing</td>
<td>38</td>
</tr>
<tr>
<td>CogState</td>
<td>39</td>
</tr>
<tr>
<td>Conclusion</td>
<td>45</td>
</tr>
<tr>
<td>Aims of Investigation</td>
<td>46</td>
</tr>
<tr>
<td>Rationale</td>
<td>46</td>
</tr>
<tr>
<td>Research Questions</td>
<td>48</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>48</td>
</tr>
<tr>
<td>Construct validation hypotheses</td>
<td>48</td>
</tr>
<tr>
<td>Criterion validation hypotheses</td>
<td>50</td>
</tr>
<tr>
<td>Methods</td>
<td>51</td>
</tr>
<tr>
<td>Participants</td>
<td>51</td>
</tr>
<tr>
<td>Procedures</td>
<td>57</td>
</tr>
<tr>
<td>Measures</td>
<td>58</td>
</tr>
<tr>
<td>Demographics and Medical Background</td>
<td>58</td>
</tr>
<tr>
<td>Psychopathology</td>
<td>58</td>
</tr>
<tr>
<td>PHQ-2</td>
<td>58</td>
</tr>
<tr>
<td>GAD-2</td>
<td>58</td>
</tr>
<tr>
<td>Sleep</td>
<td>59</td>
</tr>
<tr>
<td>SPQ</td>
<td>59</td>
</tr>
<tr>
<td>Subject Area</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Substance Abuse</td>
<td>59</td>
</tr>
<tr>
<td>The CAGE Questionnaire</td>
<td>59</td>
</tr>
<tr>
<td>Premorbid Intelligence</td>
<td>59</td>
</tr>
<tr>
<td>North American Adult Reading Test (NAART)</td>
<td>59</td>
</tr>
<tr>
<td>Neuropsychological Functioning</td>
<td>60</td>
</tr>
<tr>
<td>CogState tasks</td>
<td>60</td>
</tr>
<tr>
<td>Processing/psychomotor speed</td>
<td>62</td>
</tr>
<tr>
<td>Visual attention</td>
<td>62</td>
</tr>
<tr>
<td>Working memory</td>
<td>63</td>
</tr>
<tr>
<td>Visual learning and memory</td>
<td>64</td>
</tr>
<tr>
<td>Executive function</td>
<td>65</td>
</tr>
<tr>
<td>Verbal learning and memory</td>
<td>66</td>
</tr>
<tr>
<td>Traditional Neuropsychological Tests</td>
<td>67</td>
</tr>
<tr>
<td>Processing/psychomotor speed</td>
<td>67</td>
</tr>
<tr>
<td>Visual attention</td>
<td>68</td>
</tr>
<tr>
<td>Working memory</td>
<td>69</td>
</tr>
<tr>
<td>Visual learning and memory</td>
<td>70</td>
</tr>
<tr>
<td>Executive function</td>
<td>70</td>
</tr>
<tr>
<td>Verbal learning and memory</td>
<td>71</td>
</tr>
<tr>
<td>Hopkins Verbal Learning Task-Revised (HVLT-R)</td>
<td>71</td>
</tr>
<tr>
<td>Data Analyses</td>
<td>72</td>
</tr>
<tr>
<td>Design</td>
<td>72</td>
</tr>
<tr>
<td>Statistical Analyses</td>
<td>72</td>
</tr>
<tr>
<td>Construct validity</td>
<td>72</td>
</tr>
</tbody>
</table>
Appendix D: Informed Consent-Partners .................................................................137
Appendix E: Demographics Questionnaire for HCT Survivors .........................139
Appendix F: Demographics Questionnaire for HCT Partners ............................143
Appendix G: PHQ-2, GAD-2, SPQ ....................................................................146
Appendix H: Test order and administration times ..............................................147
Appendix I: CogState Tutorial Script .................................................................148
Appendix J: Debriefing Sheet ............................................................................149
Appendix K: Recruitment Flyer .........................................................................150
Appendix L: Personal communication ...............................................................151
Appendix M: Personal communication ...............................................................153
Appendix N: Normative data for Adapted CogState Brief Battery tasks ..........154
Appendix O: Effect sizes and 95% confidence intervals for correlations between
CogState and traditional tests .........................................................................155
List of Tables

Table 1. Neuropsychological tests, abbreviations, and domains each measure.............8
Table 2. CogState and traditional tests for comparison, score directions for good performance and expected correlations ......................................................................................49
Table 3. Summary of demographic variables ..................................................................54
Table 4. Survivor diagnoses, treatment, regimen, and GVHD prevalence........................56
Table 5. CogState tests and traditional tests compared....................................................73
Table 6. Psychosocial functioning of spousal partners and HCT survivors ..................75
Table 7. Traditional test means, standard deviations, and percentages of HCT survivor and spousal partner impairment ..........................................................................................76
Table 8. CogState task means, standard deviations, and percentages of HCT survivor and spousal partner impairment ..........................................................................................77
Table 9. Pearson’s product-moment correlations for CogState and traditional neuropsychological measures in HCT survivors .................................................................78
Table 10. Multitrait multi-method matrix of Pearson’s product-moment correlations for HCT survivors ..........................................................................................................................81
Table 11. Summary of statistically significant Pearson’s product-moment correlations, transformed into effect sizes with confidence intervals......................................................83
Table 12. Summary of Pearson’s product-moment correlations ≥ 0.30, transformed into effect sizes with confidence intervals .................................................................85
Table 13. Summary of analysis of covariance analyses for CogState Tasks and NAART.................................................................................................................................................................................. 87

Table 14. Descriptive statistics and %NOL for HCT survivors and spousal partners.......88

Table 15. T-test results and %NOL for HCT survivors and spousal partners .................89

Table 16. T-test results and %NOL for HCT survivors and spousal partners, according to domain.................................................................................................................................................................................................91

Table 17. Differences on CogState tasks for paired HCT survivors and spousal partners ..................................................................................................................................................................................92

List of Figures

Figure 1. Effect sizes and their 95% confidence intervals for significant correlations between CogState tasks and traditional neuropsychological tests ..............................................84

Figure 2. Effect sizes and their 95% confidence intervals for correlations between CogState tasks and traditional neuropsychological tests for $r’s \geq 0.3$.................................86
Introduction

Hematopoietic cell transplantation (HCT), also known as bone marrow transplantation or hematopoietic stem cell transplantation, is a life-saving procedure for patients with a host of medical conditions. Involving both high-dose chemotherapy and total body irradiation as a part of the preparative regimen, HCT encompasses a multitude of risks, both physical and psychological, including cognitive deficits. An important patient-reported component of quality of life (Ahles & Saykin, 2001), cognitive functioning also has important implications for providing informed consent, given risks associated with the procedure and the strict post-transplant care regimen. Studies have identified cognitive deficits in HCT candidates, both prior to transplant and through recovery, largely within executive functioning, memory, attention, and psychomotor speed (Beglinger et al., 2007; Jacobs, Small, Booth-Jones, Jacobsen & Fields, 2007; Syrjala, Dikmen, Langer, Roth-Roemer, & Abrams, 2004). Although HCT-associated cognitive decline occurring within one year of transplantation is well documented (Ahles, Tope, Furstenberg, Hann, & Mills, 1996; Chang, Meadows, Orav, & Antin, 2009; Harder et al., 2005; Syrjala et al., 2004), to date only two studies have investigated cognitive function in HCT survivors beyond one year (Meadows et al., 2013; Syrjala et al., 2011). Further, studies demonstrating neuropsychological decline have made use of numerous measures and varying impairment criteria, making results difficult to quantify and compiling findings across studies challenging. Another difficulty with the current literature base is the use of traditional neuropsychological testing. Although traditional neuropsychological tests are generally regarded as the “gold standard” of assessment, frequent use in the HCT population may be less than optimal. Many of the tests used are susceptible to practice effects and thus may be inappropriate for repeated assessment in longitudinal studies. Further, most traditional neuropsychological test
batteries must be administered by a neuropsychologist or highly trained psychometrician, adding cost to the already difficult task of longitudinal investigation.

The CogState brief battery, a recently developed computerized test battery, appears to be a means of addressing these concerns. This battery, which takes approximately 45 minutes to complete, is composed of a subset of the tasks from the full CogState battery. It has been validated for use with number of cognitively sensitive medical populations (Maruff et al., 2009) and has been shown to be particularly useful for patients minimizing practice effects over repeated uses. This aspect is particularly advantageous, given the limitations of using traditional neuropsychological measures to evaluate cognitive in HCT candidates and survivors over time.

The following literature review will provide an overview of HCT and risks associated with the procedure, both physical and psychological. Documented neuropsychological impairment and methods of measurement will likewise be reviewed and discussed. Finally, rationale and methods for the present study, which investigated the validity of the Adapted CogState Brief Battery, will be described. Study findings will be discussed in the context of prior research, which investigated the validity of the CogState brief battery as well as the HCT literature as a whole.

**HCT Overview**

HCT is an intense, life-saving procedure involving the transplantation of the tissue, which drives blood cell formation, including white cells to fight infection and provide immunity, platelets to support clotting and red cells to carry oxygen to the body tissues.

There are two primary types of HCT: autologous and allogeneic. Autologous transplantation involves the use of a patient’s own disease-free cells collected from an earlier point in time. Allogeneic transplantation, on the other hand, involves the transplantation of
marrow obtained from a human leukocyte antigen-matched donor or from cord blood (Bishop, Welsh, Coons, & Wingard, 2001; Bone Marrow Foundation, 2001).

**Pre-HCT Procedures**

Prior to transplantation, an HCT candidate typically undergoes extensive physical and psychological evaluation and receives formal education about the HCT process. A primary caregiver, to be available for post-transplant care, is identified and also educated. Education may include both one-on-one instruction and group classes, consisting of verbal and written information.

After finalizing the decision to undergo transplantation and identifying a donor (self or other, depending upon the graft source), the HCT candidate is hospitalized for preparative conditioning. During this time, the patient undergoes high-dose chemotherapy and possibly total body irradiation, for the purposes of eradicating cancer cells and providing immunosuppression to prevent rejection of the transplanted donor graft. A number of physical side effects are associated with conditioning, including liver damage, acute renal failure, pneumonitis, and hemorrhagic cystitis, in addition to mouth sores, nausea, vomiting, diarrhea, skin changes and hair loss. Further, conditioning-associated toxicity renders patients immunologically incompetent; thus, they must remain in protective, germ-free isolation throughout (Bishop et al., 2001).

**HCT Transplantation**

The actual transplantation procedure is quite uncomplicated, similar in process to a blood transfusion. During the two weeks that follow the transplant, however, HCT patients require the most intense care. As the new bone marrow has minimal function, patients continue to be immunocompromised. Patients must thus be protected from infection through the use of
prophylactic antibiotics and antimicrobials and/or blood transfusions and daily nutritional support, remaining hospitalized until engraftment occurs or the bone marrow recovers and begins to function again, typically 10-28 days post-transplant. Discharge from the hospital is dependent on the patient’s overall medical functioning, as measured by blood and marrow counts, absence of fever or infection, and presence of oral intake, control of nausea, diarrhea, and vomiting. As such, some patients must remain hospitalized for up to a month.

Post-HCT Physical Complications

Patients may experience a myriad of deleterious side effects immediately following transplantation and throughout the long recovery process, many of which persist for years. Shortly after the procedure, patients may experience serious infection, hemorrhage, or organ toxicity. For allogeneic patients, the bone marrow may not engraft and new bone marrow may be destroyed; more commonly, graft versus host disease (GVHD) may develop. GVHD is the most major complication of HCT; it refers to the situation in which grafted stem cells attack the body’s systems. There are two primary forms: acute GVHD, which occurs within the first 100 days of HCT, and chronic, occurring at or after 100 days post-transplant. Acute GVHD, which affects roughly 40% of HLA-matched allogeneic recipients (Ferrara, Levine, Reddy & Holler, 2009), may develop into chronic GVHD, though it is also possible for chronic GVHD to occur on its own, at 100 or more days post-transplant.

GVHD primarily affects the liver, skin, lungs, or gastrointestinal tract. Skin-related GVHD complications are characterized by rash and/or blistering, while GVHD gastrointestinal complications may include ulcerations, diarrhea, anorexia, vomiting, and bleeding. GVHD liver complications are encompassed by endothelialitis, lymphocytic infiltration of the portal areas, pericholangitis, and bile-duct destruction.
Patients who develop GVHD require ongoing treatment with immunosuppressive drugs such as cyclosporine, corticosteroids, tacrolimus, mycophenolate mofetil (MMF), and methotrexate, further increasing their risk for infection or other complications.

HCT patients also face the risk of long-term complications affecting the endocrine system (e.g. hypothyroidism or hypoadrenalism), eyes (e.g., cataracts, keratoconjunctivitis sicca, or microvascular retinopathy), skeleton (osteopenia or avascular necrosis), kidneys and bladder, (e.g. nephropathy or bladder dysfunction), and neurological and vascular systems. Other late complications of HCT include respiratory problems, recurrent infections, eye and mouth dryness, cataracts, hip problems, hormonal disturbances, reduced stamina, sleep disturbance, second malignancies, fatigue, and sexual difficulties.

**Post-HCT Psychological Complications**

Given the extensive physical risks associated with HCT, the likewise large psychological impact of the procedure is not surprising. From the outset, when considering transplantation, HCT candidates are faced with a barrage of potential stresses. Some patients may have a difficult time opting to undergo the procedure, given that their cancer is in remission (Bishop et al., 2001). Other potential stressors, identified by Lesko (1989), may include medical assessment, tissue typing procedures, relocation, familial role shifts, vocational changes, and relapse prevention procedures prior to HCT.

Research suggests that 15-55% of HCT patients experience psychological distress at some point throughout the transplant trajectory (Lee et al., 2005; Prieto, Blanch, Atala, Carreras, Rovira, Cirera, & Gastó, 2002; Wolcott, Wellisch, Fawzy & Landsverk, 1986). HCT candidates and survivors alike report a variety of psychological symptoms, including impaired social function and emotional distress (Wettergren, Langius, Björkholm, & Björvell, 1997).
Clinically, anxiety and depression appear to occur most commonly and have been documented in HCT candidates prior to admission. Clinical anxiety alone or combined with depression has been identified in at least 30% of patient samples awaiting HCT (Jenkins, Linnington, & Whittaker, 1991; Wettergren et al., 1997). Prieto et al. (2005) identified depression in 11.4% of subjects at admission, anxiety in 22.7%, and anxiety and/or depression in 26.8%, whereas Lee et al. (2005) reported pre-transplant anxiety and/or depression in 55% of patients. Notably, this broad range of estimation may be in part due to differences in psychopathology measurement. Other patient characteristics may be associated with presentation; some studies suggest a higher prevalence of depression in females (DeMarinis, Barsky, Antin, & Chang, 2008) and in females with GHVD (Syrjala et al., 2004).

Further differences in the prevalence of anxiety and depression, according to time point in the transplant trajectory, have also emerged. Anxiety appears to be most intense for patients at the time they are admitted to the hospital, whereas depression has been shown to increase following admission. Some work has demonstrated that both begin to subside two weeks after transplantation (Prieto et al., 2005).

However, other work suggests that psychological distress increases before subsiding approximately 100 days after transplantation (Lee et al., 2005; Syrjala et al., 2004). Several studies have indicated improvement beyond baseline at one-year post transplant (Chang, Orav, McNamara, Tong, & Antin, 2005; Hjermstad et al., 1999; Lee et al., 2005; Syrjala et al., 2004). In general, however, it appears that some psychological morbidity remains after patients physically recover from HCT; Syrjala and colleagues (2004) suggested that full psychological recovery occurs between three and five years post-transplant.
Neuropsychological Risks

Recently, neuropsychological effects of HCT have received increased attention in the literature (Andrykowski et al., 1990, 1992; Booth-Jones, Jacobsen, Ransom, & Soety, 2005; Harder et al., 2002, 2005; Harder, Duivenvoorden, Van Gool, & VanDen Bent, 2006, Jacobs, Small, & Booth-Jones, 2007a; Meyers et al., 1994, Schultz-Kindermann et al., 2007; Syrjala, Dikmen, Langer, Roth-Roemer, & Abrams, 2004a). An important aspect of quality of life, neuropsychological functioning may also have important implications for the lengthy post-HCT recovery period, as decreased cognitive capacity could inhibit survivor adherence to stringent post-HCT measures. While there is a gap in the literature surrounding neuropsychological functioning and medical adherence in HCT patients, cognitive deficits were recently found to be predictive of poor medical adherence in menopausal women and diabetic and hyperlipidemia patients (Stilley, Bender, Dunbar-Jacob, Sereika, & Ryan, 2010). Neuropsychological deficits have likewise been linked to decreased adherence in HIV-infected adults, particularly with complex regimens (Hinkin et al., 2002). Other work has highlighted the importance of normative cognitive functioning in medical management. A study of schizophrenic outpatients found adequate memory and conceptualization functioning to be predictive of patient ability to manage medications, over and above age, gender, education, symptom severity, and attitude surrounding medication (Jeste et al., 2003).

Studies examining neuropsychological deficits in HCT candidates and survivors have done so using cross-sectional and prospective study designs. A handful of studies have examined patient cognitive functioning longitudinally and to date, one has used a prospective, within-subjects design. A selective review follows, restricted to studies that have reported sample percentages of impairment. Trends in the literature will be highlighted in addition to limitations.
associated with each study design. Neuropsychological tests used within studies included in this literature review will be referenced according to their abbreviated name. See Table 1 for a listing of these referenced tests, including the full measure name, its abbreviated acronym, and domains each is purported to measure.

Table 1

Neuropsychological tests, abbreviations, and domains each measures

<table>
<thead>
<tr>
<th>Test name</th>
<th>Test abbreviation</th>
<th>Domains measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alertness subtest, TAP</td>
<td>TAP alertness</td>
<td>Simple reaction time, phasic alertness</td>
</tr>
<tr>
<td>Benton Visual Retention Test</td>
<td>BVRT</td>
<td>Learning, memory, visual memory, visual perception, and visuoconstructive abilities</td>
</tr>
<tr>
<td>Block design</td>
<td>...</td>
<td>Visuospatial and constructional ability</td>
</tr>
<tr>
<td>Buschke Selective Reminding</td>
<td>...</td>
<td>Memory, anterograde learning</td>
</tr>
<tr>
<td>California Verbal Learning Test I, II</td>
<td>CVLT, CVLT-II</td>
<td>Verbal memory, learning, memory</td>
</tr>
<tr>
<td>Category Fluency subtest, RWT</td>
<td>...</td>
<td>Divergent thinking</td>
</tr>
<tr>
<td>Category Word Fluency</td>
<td>...</td>
<td>Attention, executive function</td>
</tr>
<tr>
<td>Cognitive Failure Questionnaire</td>
<td>CFQ</td>
<td>Subjective cognitive functioning</td>
</tr>
<tr>
<td>Connors’ Continuous Performance Test</td>
<td>CPT-II</td>
<td>Attention</td>
</tr>
<tr>
<td>Controlled Oral Word Association subtest, MAE</td>
<td>COWA</td>
<td>Simple executive function, verbal fluency, memory, executive function, executive control</td>
</tr>
<tr>
<td>The d2 Cancellation Test</td>
<td>d2 cancellation</td>
<td>Selective attention</td>
</tr>
<tr>
<td>The d2 Test of Attention</td>
<td>d2</td>
<td>Attention, executive function</td>
</tr>
<tr>
<td>Delirium Rating Scale</td>
<td>DRS</td>
<td>Delirium</td>
</tr>
<tr>
<td>Dementia Rating Scale</td>
<td>...</td>
<td>Attention, initiation/perseveration (frontal lobe executions), construction, conceptualization, memory</td>
</tr>
<tr>
<td>The European Organization for Research and Treatment of Cancer, Quality of Life Questionnaire-C30</td>
<td>EORTC QLQ-C30</td>
<td>Quality of life, subjective cognitive functioning</td>
</tr>
<tr>
<td>Test</td>
<td>Short Form</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------------------------------------</td>
<td>------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Finger Oscillation Test</td>
<td>...</td>
<td>Motor function</td>
</tr>
<tr>
<td>Finger Tapping Test</td>
<td>FTT</td>
<td>Psychomotor functions, psychomotor speed</td>
</tr>
<tr>
<td>Groninger Intelligence Test</td>
<td>...</td>
<td>General intelligence, complex tasks</td>
</tr>
<tr>
<td>Grooved Pegboard Dominant and Nondominant</td>
<td>GPB, GPB-D, GPB-ND</td>
<td>Psychomotor speed, motor function, motor strength, speed and dexterity, motor speed, coordinated and fine motor speed</td>
</tr>
<tr>
<td>Hand Dynamometer</td>
<td>...</td>
<td>Motor strength, speed, and dexterity</td>
</tr>
<tr>
<td>Hopkins Verbal Learning Test, Hopkins Verbal Learning Test-Revised</td>
<td>HVLT, HVLT-R</td>
<td>Verbal fluency, verbal learning, memory</td>
</tr>
<tr>
<td>Leistungs-Pruf-System (Achievement Measure System)</td>
<td>LPS</td>
<td>Reasoning</td>
</tr>
<tr>
<td>Letter fluency subtest, RWT</td>
<td>...</td>
<td>Divergent thinking</td>
</tr>
<tr>
<td>Memorial Delirium Assessment Scale</td>
<td>...</td>
<td>Delirium</td>
</tr>
<tr>
<td>Mini Mental Status Exam</td>
<td>MMSE</td>
<td>Global cognitive function</td>
</tr>
<tr>
<td>Modified MMSE</td>
<td>3MS</td>
<td>Global cognitive function</td>
</tr>
<tr>
<td>Multilingual Aphasia Examination</td>
<td>MAE</td>
<td>Aphasic disorder</td>
</tr>
<tr>
<td>Multiple Abilities Questionnaire</td>
<td>MAQ</td>
<td>Subjective cognitive function</td>
</tr>
<tr>
<td>National Adult Reading Test</td>
<td>NART</td>
<td>Intelligence, complex tasks</td>
</tr>
<tr>
<td>Reaction Time Task</td>
<td>...</td>
<td>Psychomotor functions, speed of information processing, psychomotor speed</td>
</tr>
<tr>
<td>Regensburger Wortflussigkeits-Test (Regensburg Word Fluency Test)</td>
<td>RWT</td>
<td>Executive functions, divergent thinking</td>
</tr>
<tr>
<td>Repeatable Battery for the Assessment of Neuropsychological Status</td>
<td>RBANS</td>
<td>Battery for dementia assessment which includes tests of attention, language, immediate memory, delayed memory visuospatial/constructional abilities</td>
</tr>
<tr>
<td>Rey Complex Figure Test</td>
<td>RCFT</td>
<td>Visual memory, learning</td>
</tr>
<tr>
<td></td>
<td>RCFT copy</td>
<td>Visuospatial and constructional ability, visuospatial organization, complex tasks, intelligence</td>
</tr>
<tr>
<td></td>
<td>RCFT recall</td>
<td>Memory</td>
</tr>
<tr>
<td></td>
<td>RCFT recognition</td>
<td>Memory, learning</td>
</tr>
<tr>
<td>Ruff two and seven test</td>
<td>...</td>
<td>Sustained visual attention</td>
</tr>
<tr>
<td>Stroop ColorWord Test</td>
<td>Stroop</td>
<td>Attention, executive function, complex executive function</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Test</th>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol Digit Modalities Test</td>
<td>SDMT</td>
<td>Visual processing, divided attention</td>
</tr>
<tr>
<td>Temporal Orientation Test subtest of the Benton Iowa Screening</td>
<td></td>
<td>Orientation to time</td>
</tr>
<tr>
<td>Battery for Mental Decline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test battery for Attentional Performance</td>
<td>TAP</td>
<td>Battery includes tests of visuo-spatial, non-spatial and executive attention, and working memory.</td>
</tr>
<tr>
<td>TAP Go/No Go</td>
<td></td>
<td>Selective attention</td>
</tr>
<tr>
<td>Trail Making Test A &amp; B</td>
<td>TMT-A</td>
<td>Attention, concentration, executive function, graphomotor speed, simple attention, speed of information processing, visual scanning, visuospatial/constructional ability, working memory</td>
</tr>
<tr>
<td>TMT-B</td>
<td></td>
<td>Complex attention/motor, attention, executive function, simple executive function, processing speed, divided attention, cognitive attention, concentration, visual scanning, visuospatial/constructional ability, working memory</td>
</tr>
<tr>
<td>Verbal Fluency Test</td>
<td></td>
<td>Attention, concentration, executive function</td>
</tr>
<tr>
<td>Verbaler Lern-und Merkfähigkeitstest (German adaptation of the</td>
<td>VLMT</td>
<td>Verbal learning, verbal long-term memory, verbal working memory, verbal recognition</td>
</tr>
<tr>
<td>Auditory-Verbal Learning Test)</td>
<td></td>
<td>Verbal recognition</td>
</tr>
<tr>
<td>Verbal recognition Trial W-F</td>
<td></td>
<td>Verbal learning</td>
</tr>
<tr>
<td>VLMT Sum Trials 1-5</td>
<td></td>
<td>Verbal learning</td>
</tr>
<tr>
<td>VLMT Trial 1</td>
<td></td>
<td>Verbal working memory</td>
</tr>
<tr>
<td>VLMT Trial 5</td>
<td></td>
<td>Verbal learning</td>
</tr>
<tr>
<td>VLMT Trials 5 minus 7</td>
<td></td>
<td>Verbal long-term memory</td>
</tr>
<tr>
<td>VLMT Trials 5 minus 6</td>
<td></td>
<td>Verbal long-term memory</td>
</tr>
<tr>
<td>Wechsler Memory Scale-III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual Memory Span</td>
<td></td>
<td>Visual working memory</td>
</tr>
<tr>
<td>Visual Reproduction</td>
<td></td>
<td>Visual memory, memory</td>
</tr>
<tr>
<td>Spatial Span (Span)</td>
<td></td>
<td>Visual scanning, working memory</td>
</tr>
<tr>
<td>Wechsler Adult Intelligence Test-III, Revised</td>
<td>WAIS-R</td>
<td>Intelligence</td>
</tr>
<tr>
<td>Digit Span</td>
<td></td>
<td>Attention, executive function, simple attention, memory, verbal working memory, information-processing speed</td>
</tr>
</tbody>
</table>
Cross-Sectional Studies

A series of studies have investigated neuropsychological functioning in HCT samples cross-sectionally, prior to or after transplantation. These studies will first be reviewed.

**Pre-HCT.** Andrykowski et al. (1992) evaluated cognitive function in 46 HCT candidates before transplantation, during the conditioning period. Participants’ functioning was assessed across four domains: (a) motor (GPB [Kløve, 1964] and Finger Oscillation Test [Halstead, 1947]), (b) memory (Buschke Selective Reminding [Buschke, 1973] and BVRT [Benton, 1974]), (c) attention (Digit Span [Wechsler, 1981], Ruff 2 and 7 test [Ruff, Evans, & White, 1986]), and (d) complex attention/motor (TMT-B [Reitan & Wolfson, 1985] and Digit Symbol [Wechsler, 1981]). Impairment was classified as “probable” in patients scoring ≥1.5 standard deviations below test norms and “definite” in those scoring ≥2 standard deviations below test norms. Analyses indicated that 56% of the sample met criteria for probable or definite impairment. Impairment was most prominent on memory tests, with 34-43% scoring in the probable impairment range (scoring ≥1.5 standard deviations below test norms). Less than 6% of patients scored in the impaired range on tests measuring attention. Performance on complex attention/motor tasks was varied: 33% of the sample scored in the impaired range on TMT-B, regarded as a highly sensitive indicator of impairment in complex cognitive processes, given that it requires cognitive flexibility in the use of visual-motor, attention, conceptual tracking, and
sequencing skills. In contrast, on the Digit Symbol task, another test of complex motor/attention, only 6% demonstrated impairment.

In the context of the normal distribution, impairment rates are high overall, given that 6.7% fall \( > 1.5 \) standard deviations below test norms.

Harder et al. (2005) also assessed cognitive performance in HCT candidates awaiting transplantation, as compared with hematologic patients. This comparison was examined in order to determine whether hematologic patients could serve as a reference group for future prospective studies of HCT-associated cognitive change. HCT candidates had received previous treatment with systemic chemotherapy and/or radiotherapy \((n = 101)\); hematological patients \((n = 82)\) had previously been treated with systemic chemotherapies and/or involved field radiotherapy. Thirty-four percent of the HCT candidates were scheduled for autologous HCT, 64% for allogeneic transplant. Assessment of HCT candidates was undertaken approximately 21 days prior to transplant. Participants were assessed in four cognitive domains: (a) memory and learning (Dutch version of the CVLT [Mulder, Dekker, & Dekker, 1996], the Rey Complex Figure Test and recognition trial [Osterrieth, 1944], the BVRT [Benton, 1974]), (b) attention and executive functions (Category Word Fluency [Ruff, Evans, & White, 1986], Digit Span [Wechsler, 1997], the abbreviated Stroop [Stroop, 1935], TMT-A, TMT-B [Reitan & Wolfson, 1985], and d2 [Brickenkamp, 1981]), (c) visuospatial and constructional ability (Rey complex figure test-copy trial [Osterrieth, 1944] and Block design [Wechsler, 1997]), and (d) psychomotor functions (Digit symbol [Wechsler, 1997], FTT [Reitan & Davison, 1974], and the Reaction Time Task [Middelkoop, Vink, & Lanser, 1996]). The Dutch version of the CFQ (Broadbent, Cooper, FitzGerald, Parkes, 1982) was administered as a measure of subjective cognitive functioning and the Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983),
Multidimensional Fatigue Inventory (Smets, Garssen, & Bonke, 1995), and EORTC-QLQ (Aaronson et al., 1993) as measures of psychosocial functioning, fatigue, and quality of life, respectively.

Neuropsychological test scores were converted to z-scores; participants scoring 1.5 standard deviations below population norms on at least 4 sub-tests were classified as impaired. A measure of overall cognitive impairment was also computed based on the percentage of impaired scores relative to completed tests.

Results indicated greatest impairment in visual memory, visuospatial and constructional ability, and psychomotor functions in both groups. Although HCT candidates were slower than hematologic patients on TMT-A \((p = 0.03)\), no other significant differences were found between the groups of impaired patients. Twelve percent of the HCT patients had impaired scores on more than 20% of tests, as compared with 8.5% of hematological patients. No significant correlations were found between patients’ subjective reports of cognitive functioning and objective neuropsychological testing, though HCT candidates demonstrated higher anxiety and reported lower cognitive, emotional, and social functioning. These findings were in line with other prior work demonstrating psychosocial distress in HCT candidates.

In sum, this study demonstrated that similar patterns of cognitive decline occur in HCT candidates and hematological patients with prior systemic and/or involved-field radiotherapy. Given that no between-group differences emerged with regard to the degree or pattern of cognitive impairment, hematological patients were determined to be an adequate reference group for future longitudinal work examining cognitive change in HCT candidates during and after transplantation. Further, in the context of the normal curve, impairment rates illustrated in this study were high, since about 7% of the normal population exhibit z-scores < -1.5.
**Six months post-HCT.** In 2005, Booth-Jones and colleagues assessed cognitive function in 65 HCT survivors, 81% of whom had undergone an allogeneic transplant. Participants’ functioning was assessed 6 months following transplant, across seven neuropsychological domains: (a) simple attention (TMT-A [Reitan & Wolfson, 1985] and Digit Span [Wechsler, 1981]), (b) verbal learning (HVLT [Brandt, 1991]), (c) verbal memory (Logical memory [Wechsler, 1997]), (d) visual memory (WMS-III), (e) simple executive function (Digit Symbol [Wechsler, 1997], TMT-B [Reitan & Wolfson, 1985] and COWA [Benton & Hamsher, 1983]), (f) complex executive function (Stroop [Stroop, 1935] and WCST [Grant & Berg, 1948]) and (g) psychomotor speed (GPB-D and GPB-ND [Kløve, 1964]), using traditional neuropsychological tests. The Multiple Abilities Questionnaire (MAQ; Seidenberg, Haltiner, & Taylor, 1994), a self-report measure, provided data on subjective cognitive functioning across six domains: Attention, Language, Remote Memory, Verbal Memory, Visual-Spatial Memory, and Visual-Spatial Perception. Raw cognitive tests scores were converted to z-scores, and tests were combined into cognitive domains scores through averaging domain test z-scores. Further, a total neuropsychological performance score (TNP) was calculated by averaging all test z-scores in the battery. Z-scores below -1 were categorized as deficient. More specifically, z-scores between -1 and -1.5 were categorized as mildly deficient, and z-scores equal or lower than -1.5 were categorized as moderately to severely deficient.

Results indicated that 37% of participants had at least mild deficits in psychomotor speed (≤ -1.00 z-scores ≥ -1.5), 15% had at least mild deficits in complex executive function, 6% had at least mild deficits in verbal learning, and 3% had at least mild deficits in verbal memory. Mild or greater deficits in attention, visual memory, and simple executive function were observed in 1.5% of the sample, suggesting that lower rates than shown in the normal population were
observed. Moderate or severe deficits ($z < -1.5$) in psychomotor speed were identified in 14% of the sample, which is higher than the normal population (i.e. the normal curve suggests that 2-7% of the population would exhibit $z$-scores $<-1.5$). Executive function-related deficits were observed in 6% of the sample, and deficits in verbal learning, verbal memory, and visual memory in 3%. Forty-nine percent of the sample were absent of deficits; in total, 28% of participants experienced moderate or severe deficiency in at least one domain, a high percentage in the context of the normal curve, as 6.7% of the population exhibit $z$-scores $\leq -1.5$.

Nevertheless, this study was limited in a number of ways. First, cross-sectional measurement limited conclusions, such that deficits identified could have been present prior to transplantation. Further, participants’ test scores were compared with population norms, rather than a matched control sample. Also, while the study included both allogeneic and autologous transplant survivors, differences between the two groups were not quantified. Given that the majority of the sample (81%) had undergone autologous transplant, the authors suggested that results of the study generalized primarily to this group. Finally, participant treatment regimen history was not obtained. Thus, specificity with regard to the cognitive impact of individual treatment history also limits interpretation of study results.

**One to four years post-HCT.** Harder et al. (2002) assessed cognitive functioning and quality of life in 40 HCT survivors, 22-82 months post-transplant. All patients had previously received total body irradiation; 87.5% had undergone allogeneic transplant. All participants completed a comprehensive battery of neuropsychological tests and an interview related to cognitive problems, quality of life, and mood measures. Neuropsychological tests were used to assess the following domains: (a) general intelligence (Groninger Intelligence Test [Snijders, Luteijn, van der Ploeg, & Vergage, 1983]), (b) premorbid intelligence (NART [Nelson, 1982]),
(c) language (Wordfluency [Snijders, 1983]), (d) memory (CVLT [Delis, Kramer, & Kaplan, 1996]), (e) visuospatial organization and visual memory (Rey Complex Figure Test-recall [Osterrieth, 1944]), (f) attention and executive function (Stroop [Stroop, 1935], Digit Span [Wechsler, 1997], TMT-A, and TMT-B [Reitan & Wolfson, 1985]), (g) psychomotor speed (FTT [Reitan & Davison, 1974]), and (h) speed of information processing (Reaction Time Task [Middelkoop et al., 1996]). EORTC-QLQ (Aaronson et al., 1993) was used to measure survivor quality of life and subjective cognitive functioning; the Profile of Mood States (McNair, Lorr, & Droppleman, 1971) was administered in order to distinguish between cognitive deficits associated with HCT treatment or psychological distress.

All neuropsychological test scores were converted into z-scores; scores falling ≥2 SDs below the standard norms were classified as impaired. A composite score, representative of overall impairment, was calculated by counting all impaired test indices. Results revealed memory impairment in 15-20.5% of patients, speed of information processing impairment in 32.5%, visuospatial organization impairment in 2.6%, and psychomotor impairment in 2.5%. Performance on the Stroop test (Stroop, 1935), a measure of attention and executive function, was impaired in 23.1% of the sample, although other test results of attention and executive function (i.e. Digit Span, TMT-A, TMT-B) did not reveal impairment. In the context of the normal curve, impairment in memory, speed of processing, and attention/executive function (as measured by the Stroop test) were quite high, as 2.3% of the normal population exhibit z-scores < 2.0; impairment rates for visuospatial organization and psychomotor speed were consistent with normal population rates.

In total, 60% of the sample scored in the impaired range on one test and 12.5% in three or more tests. Cognitive impairment was related to problems subjectively reported on the EORTC
QLQ (Aaronson et al., 1993) but not to information gathered via interview. Higher educational level and fatigue and global health, as measured by the EORTC-QLQ (Aaronson, 1993), were identified as the strongest predictors of the impairment.

Although this study was one of the earliest to evaluate long-term cognitive side effects associated with HCT, it was limited by a small sample size. Therefore, the range of survivor chemotherapeutic agents/dosages could not be assessed for their impact on functioning. Further, the retrospective design and lack of baseline assessment precluded conclusions about the cognitive impact of HCT.

Taken together, these cross-sectional studies illustrate impairment occurring at varying stages across the transplant trajectory. Before transplant, cognitive deficits in memory, psychomotor functioning, attention, and visuospatial ability were identified (Andrykowski et al., 1992; Harder et al., 2002). At six months post-transplant, deficits in memory, psychomotor functioning, attention, executive function, and verbal learning were identified (Booth-Jones et al., 2005). At two–four years post-transplant, deficits in memory, psychomotor functioning, attention, executive functioning, speed of information processing, and visuospatial ability were identified (Harder et al., 2002). Impairment rates vary, but in general, studies illustrate higher impairment than in the normal population, in the context of the normal curve.

Given that subjects in these studies were tested at one time point only, these studies are informative; however, the lack of baseline data limits interpretation of HCT survivor study findings (Booth-Jones et al., 2005; Harder et al., 2002). Collectively, the cross-sectional methods used in this study set provide limited insight into the HCT neuropsychological symptom trajectory.
Prospective Studies

Prospective studies, on the other hand, have focused upon how cognitive changes progress as HCT survivors move through varying stages of recovery. These studies will be reviewed in the following sections. Limitations and strengths of each will likewise be highlighted.

Pre-HCT through 3 months post-HCT. Schultz-Kindermann et al. (2007) investigated allogeneic HCT functioning prior to transplant ($N = 39$) and three months post-transplant ($n = 19$) across three cognitive domains: (a) attention (TMT-B [Reitan & Wolfson, 1985], TAP Go/NoGo [Zimmermann & Fimm, 2002], d2 test [Brickenkamp, 1981], TMT-A [Reitan & Wolfson, 1985], and TAP alertness [Zimmermann & Fimm, 2002]), (b) memory (Digit Span [Wechsler, 1987], VLMT Trial 1 [Helmstaedter, Lendt, & Lux, 2001], VLMT, Trial 5, and Sum Trials 1-5 [Helmstaedter et al., 2001], VLMT Trials 5 minus 6 and Trials 5 minus 7 [Helmstaedter et al., 2001], VLMT: Recognition Trial W-F [Helmstaedter et al., 2001], Visual Memory Span [Wechsler, 1987]), and (c) executive function (RWT [Aschenbrenner, Tucha, & Lange, 2000] and LPS [Horn, 1983]). All test scores were transformed to $z$-scores; in accordance with Spreen and Strauss’ recommendations (1998), $z$-scores $\leq -1.4$ were considered indicative of impairment. A composite score of all impaired scores was also calculated, based on the number of noted impaired functions. A composite score of $> 3$, reflecting impairment on three tests, was considered to be indicative of global impairment. Results were compared with normative data. Cognitive impairment was noted in all three domains at pre-transplant; 26.3% of patients demonstrated global impairment at both time points. Minimal differences were noted at the three-month follow-up, although simple reaction time worsened significantly. Although these findings paralleled previous work (Andrykowski et al., 1992; Booth-Jones, 2005), which
suggested increased deficits at 100 days post-transplant, other prospective work has demonstrated improved neuropsychological performance in HCT survivors 100 days post-transplant (Meyers, 1994; Sostak, 2003).

**Pre-HCT through 8 months post-HCT.** Meyers et al. (1994) prospectively investigated cognitive functioning in 61 HCT candidates, 41% of whom were scheduled to undergo allogeneic transplantation. Patients underwent assessment prior to hospitalization, two weeks following transplant, at discharge, and finally at 8 months post-transplant. Cognitive functioning was assessed via the Dementia Rating Scale (Mattis, 1988), which provided measures of attention, initiation/perseveration (frontal lobe executive functions), construction, conceptualization, and memory. Standardized scores $\geq 2$ standard deviations below the mean were considered to be indicative of impairment.

Impairment in executive function and memory were identified at each time point, but deficits in attention only at discharge. Memory deficits were observed to increase over the course of transplant hospitalization. Specifically, 11% demonstrated deficits in memory at baseline; at two weeks post, 26%; at discharge, 15%, and at 8 months post, 19%. Impairment in executive function was shown to decrease over time, with 41% showing deficits at baseline, 33% at two weeks post, 32% at discharge, and 29% at 8 months post. Impairment in attention was identified in 2% of participants at discharge. According to the normal curve distribution, approximately 2% of the population would exhibit $z$-scores $<-2.0$. As such, in comparison with the normal population, rates of impairment evidenced were particularly high in memory and executive function at all-time points, whereas exhibited impairment in attention was in line with what is expected in the normal population.
Of note, although 61 subjects completed baseline assessment, because of attrition due to subject unavailability, death, and disease progression, only 54 subjects completed assessments at two weeks post, 47 at discharge, and 21 at 8 months post. As such, study authors recommended caution in interpreting results, in recognition that selection bias may have favored patients with good outcomes.

**Pre-HCT through 1 year post-HCT.** Harder et al. (2004) applied methods from a previous study, prospectively investigating neuropsychological functioning in 21 HCT candidates, 76% of whom were scheduled to undergo allogeneic transplantation. Participants underwent assessment prior to transplantation, 6 months post, and one year post-HCT. They were assessed on four cognitive domains: (a) intelligence and complex tasks (Groninger Intelligence Test [Snijders et al., 1983], NART [Nelson, 1982], and the Rey Complex Figure Test-copy [Osterrieth, 1944]), (b) memory (Digit Span of the Wechsler Adult Intelligence Scale [Wechsler, 1997], the Dutch version of the CVLT [Mulder et al., 1996], and immediate recall of the Rey Complex Figure Test [Osterrieth, 1944]), (c) attention and executive functions (Stroop [Stroop, 1935], TMT-A, TMT-B [Reitan & Wolfson, 1985], and Word fluency Test), and (d) psychomotor functions and speed (FTT [Reitan & Davison, 1974] and Reaction Time Task [Middelkoop et al., 1996]). Participants were also interviewed by a neuropsychologist about subjective cognitive complaints they had experienced. Raw scores for each traditional test were converted to standardized z-scores; scores ≤ 2 SD below population norms or below the 10th percentile were classified as impaired.

Baseline testing, completed in 25 participants, revealed impairment in three domains: 4% exhibited impaired scores in intelligence, 24% in attention and executive functions, and 24% in psychomotor functions and speed. The majority of patients (80%) reported no subjective
cognitive complaints. At 6 months post-HCT, 12 participants completed testing; 25% exhibited impairment in attention and executive functions and 8% in psychomotor functions and speed; per subjective self-report, 50% reported subjective cognitive difficulty with memory and attention. At 12 months post-transplant, 11% exhibited impairment in memory, 22% in attention and executive functions, and 22% in psychomotor functions and speed; 30% reported subjective cognitive difficulty with memory.

Results of this study overall suggested that memory may slightly improve over time, but, as previously noted, attrition was problematic. Notably, subjective cognitive complaints were largely uncorrelated with actual impairment, except within subjective concentration problems noted at the one-year follow-up ($p < .004$), a finding that could be representative of patients’ awareness of cognitive change or, alternatively, a deficiency in neuropsychological test sensitivity to detect subtle cognitive change.

Notably, rates of impairment in attention and executive functions were, at all time points, high in the context of the normal curve given that 9.7% of the population exhibit z-scores ≤ 1.3. Results overall should be interpreted with caution, given the decreased sample size associated with attrition. The authors suggested that future prospective longitudinal assessments include at least 100 subjects in order to account for probable 50% attrition due to morbidity (e.g., post-transplantation complications requiring prolonged hospitalization) as well as mortality.

In a much larger prospective study of 142 allogeneic HCT candidates, Syrjala et al. (2004) assessed cognitive functioning before transplant, 80 days post, and one year post transplant. Participants completed an assessment battery at each time point, which included measures of (a) intelligence (Information [Wechsler, 1987]), (b) motor strength, speed, and dexterity (Hand Dynamometer [Reitan & Wolfson, 1985] and GPB [Kløve, 1964]), (c) attention
and processing speed (TMT-B [Reitan & Wolfson, 1985] and Digit Symbol [Wechsler, 1997]),
and (d) verbal fluency and memory (COWA [Benton & Hamsher, 1983] and HVL-T-R [Benedict, Schretlen, Groninger, & Brandt, 1998]). At the final assessment point (one year post-transplant),
subjects completed the WCST (Grant & Berg, 1948), a measure of executive function. Scores
were converted to standardized t-scores; T-scores ≤ 1 SD below the mean were classified as
impaired. Of the 142 participants enrolled in the study, all were tested at least once; 85% were
assessed before transplant, 38% at all three time points, 27% twice, and 35% only once.

Deficits in all cognitive domains were noted at each time point. Results revealed
cognitive impairment in 15-32% of the sample at pre-transplant assessment. Nearly all scores,
particularly in motor skills and verbal fluency, declined through 80 days post-transplant before
significantly improving at one year post-transplant. At one year post-transplant, most patients
returned to baseline functioning in all domains except motor dexterity function, which remained
impaired for 46% of the sample. In the context of the normal curve, this is high, since 16% of the
normal population exhibit z-scores < 1.0. The authors noted that results may be limited due to
practice effects associated with the neuropsychological tests administered, which could have
inflated test scores or masked impairments.

In another study with a large sample size, Jacobs, Small, Booth-Jones, Jacobsen, and
Fields (2007) investigated neuropsychological functioning in HCT candidates (N = 388), 79% of
whom were awaiting autologous transplantation. Participants were assessed at three key points:
prior to HCT, 6 months post, and 12 months post. A sequential longitudinal study design was
used in order to examine whether cognitive functioning improved after HCT, independent of
practice effects. Thus, participants were randomized to testing at all three time points, at 6 and 12
months post, or only at 12 months post-transplant. Participants underwent assessment of four
domains: (a) attention (CPT-II [Conners, 2000]), (b) memory (California Verbal Learning Test-II and Logical Memory and Visual Reproduction [Wechsler, 1997]), (c) executive functioning (Digit Symbol [Wechsler, 1997], TMT-A and B [Reitan & Wolfson, 1985], COWA [Benton & Hamshner, 1983] and Stroop [Stroop, 1935]), and (d) motor speed (GPB [Kløve, 1964]). Raw scores for each test were converted to $z$-scores and categorized according to domain measurement. Participant performance was classified at two levels of impairment: if either the mean of the $z$-scores in a domain or the mean of all the $z$-scores was $\leq -1.5$ or $\leq -1.0$.

Results revealed impairment in all domains prior to transplant; 33.22% of the sample demonstrated impairment in motor speed prior to transplant and 2.85-8.39% of the sample in other domains. At one year post-transplant, functioning in all domains except attention had significantly improved, in accordance with Syrjala and colleagues’ findings (2004). The study suggested that pre-HCT deficits remitted to within normal limits by one year post-transplant.

However, it is also possible that score improvement was an artifact of practice effects, as repeated measurement was associated with higher motor speed performance and total neuropsychological performance score. Furthermore, study attrition, due to death, illness, disinterest, and participant unreachability, was nearly 70%. Consequently, the sample may have been representative of a healthier group of HCT patients, inflating the degree to which subjects demonstrated cognitive improvement. Regardless, the degree to which deficits were shown in motor speed was striking, considering that in the context of the normal curve, 6.7% of the normal population exhibit $z$-scores $\leq 1.5$.

More recently, Scherwath and colleagues (2012) undertook a prospective investigation of medical correlates associated with cognitive decline across the HCT trajectory. Participants ($N = 102$) were evaluated before transplant, approximately 100 days post, and approximately one year
post, across four domains: (a) attention (Tests of Attentional Performance [Zimmerman & Fimm, 2002]), (b) memory (Wechsler Memory Scale-Revised digit span and visual span [Harting et al., 2000] and the Verbal Learning and Memory Test [Helmstaedter et al., 2001]), (c) executive function (Regensburg Word Fluency Test [Aschmbrenner et al., 2000]), and (d) motor function (Grooved Pegboard Test [Trites, 1977]). Raw scores were converted into $z$-scores and restricted to a range from -3.00 to +3.00 to correct for outliers. Scores of $z \leq 1.4$ were considered impaired.

Results indicated that prior to transplant, 24% of participants showed impairment, most prominently in verbal fluency. At 100 days post-transplant, 16% of participants showed impairment, most prominently in alertness and verbal delayed recall. At one year post-transplant, up to 19% evidenced impairment on fine motor function and verbal delayed recall. These rates of impairment are notable in the context of the normal curve, as 8% of the population exhibit scores in this range.

Overall, the authors found that participants performed consistently below test norms in verbal fluency and retrieval, alertness, divided attention, and fine motor function, in support of others’ research (Syrjala et al. 2004; Beglinger et al., 2007; Jacobs et al., 2007). Few associations were noted between medical variables and cognitive function, although higher participant age was associated with worse verbal memory. Not surprisingly, higher age and premorbid IQ were associated with better word fluency. In general, findings supported others’ findings, but results were limited by a lack of a matched control group, in order to control for both practice effects and cognitive changes associated with age.

In the only study to use a within-subjects design, Friedman et al. (2009) prospectively evaluated cognitive change in HCT candidates. Of those participating ($N = 117$), 48% underwent allogeneic transplantation. Subsequent neuropsychological assessments were completed
following transplant. Participants were assessed on four cognitive domains: (a) learning (HVLT [Brandt, 1991]), (b) attention (TMT-A [Reitan & Wolfson, 1985] and Digit Span [Wechsler, 1997]), (c) executive function (TMT-B [Reitan & Wolfson, 1985] and COWA [Benton & Hamsher, 1983]) and (d) processing speed (Digit Symbol [Wechsler, 1997]). Raw scores were converted to standardized scores, and participants were classified as impaired if they obtained z-scores of \leq -1.5 on at least two tests or a z-score of \leq -2.0 on at least one test. Results indicated that 39% of HCT candidates were impaired at baseline. At each time point, impairment was higher than that found in the normal population. At 6 weeks post-transplant, 33 participants were reassessed, and at 28 weeks post, 32 were reassessed. In comparison to baseline performance, 47% of participants displayed reliable decline on at least one test when they were reassessed at 6 weeks post-transplant. Between baseline and 6 weeks post-HCT, 25% of participants demonstrated decline on the HVLT and 9% on TMT-B and Digit Symbol; 47% of patients (n=15) showed declines at 6 weeks post-transplant. Of the 15 patients who showed declines at 6 weeks post-transplant, 33% (n=5) demonstrated further decline at 28 weeks post-transplant, whereas 7 improved and 2 showed no change. Of the participants who did not show a decline at 6 weeks post-transplant (n=17), 3 showed a decline in performance at the 28-week follow-up point. Overall, baseline impairment rates were high in the context of the normal curve: 6.7% of the normal population exhibit z-scores \leq -1.5. Declines exhibited over time were likewise much higher than scores generally shown in normal population.

Although this within-subjects design provided measurement of individual cognitive change over time, findings are limited by the absence of an external control group, which could clarify the role of disease or treatment variables upon patient functioning. Furthermore, findings
are limited by the use of repeated measures, as potential practice effects may have masked cognitive deficits.

**Pre-HCT through 18 months post-HCT.** In an investigation of predictors of neuropsychological change in patients with chronic myelogenous leukemia and myelodysplastic syndrome, Meadows and colleagues (2013) evaluated cognitive performance in patients who received either allogeneic HCT or standard treatment (e.g. chemotherapy and supportive care). A total of 106 HCT candidates were recruited for study participation. Of these participants, 91 were patients with chronic myelogenous leukemia and 15 had myelodysplastic syndrome. Participants were evaluated baseline, at 12 months, and at 18 months post-treatment (e.g. HCT or standard). They were assessed across four neuropsychological domains, including (a) attention (TMT-A [Reitan & Wolfson, 1985]), (b) executive functioning (TMT-B [Reitan & Wolfson, 1985], Verbal Fluency Test [Benton and Hamsher, 2002], Stroop Color-Word Test [Golden and Freshwater, 2002]), (c) anterograde learning and memory (Buschke Selective Reminding Test [Masur et al., 1989]), and (d) processing speed (GPB [Klove, 1963]). Raw scores for each test were converted to $z$-scores, and a composite score for each domain was calculated by averaging test scores for each domain. $Z$-scores ≤ -1.4 were classified as impaired. In the context of the normal curve, $z$-scores < -1.4 are exhibited by 8% of the population.

At baseline, 23% of the entire sample illustrated impaired performance on at least one test, 18% on two tests, and 39% on three or more tests; 19% scored within normal limits. At 18 months post-treatment, 36% were functioning within normal limits; 15% illustrated impaired performance on one test, 19% on two tests, and 31% on more than two tests.

In general, scores improved over time; notably, improvement was statistically significant for Executive Function and Memory domains. Findings indicated that treatment type (including
HCT) was not a significant predictor of change within one year of treatment. Rather, age, estimated IQ, accelerated chronic myelogenous leukemia and myelodysplastic syndrome emerged as multivariate predictors of cognitive change at 12 months post-treatment compared with baseline. Memory and attention at baseline, age, estimated IQ, and a mental component score from the Medical Outcomes 26 Item Short Form (SF-36; McHorney, Ware, & Raczek, 1993) emerged as multivariate predictors of change between 18 months and baseline. In general, findings supported prior work identifying cognitive impairment pre-transplant, as well as cognitive recovery over time. In contrast to others’ work, it did not show HCT to be a predictor of cognitive decline.

**Pre-HCT through five years post-HCT.** To date, one longitudinal investigation has evaluated cognitive functioning in HCT survivors beyond one year. In 2011, Syrjala and colleagues undertook a prospective study, evaluating HCT candidates ($n = 161$) alongside case-matched controls ($n = 66$) 2-14 days before and approximately 80 days, one year, and five years post-transplant. Participants underwent assessment across the following cognitive domains: (a) executive control (COWA [Benton & Hamsher, 1983]), (b) information-processing speed (Digit symbol [Wechsler, 1997]), (c) verbal list-learning (HVLT-R [Benedict, Schretlen, Groninger, & Brandt, 1998]), (d) motor speed and dexterity (GPB [Kløve, 1964]), (e) visual scanning, graphomotor speed, and attention (TMT-A [Reitan & Wolfson, 1985]), and (f) executive function (TMT-B [Reitan & Wolfson, 1985]).

$T$ scores $< 40$ (e.g. more than one standard deviation below the mean) were characterized as impaired. Notably, a $T$-score of 40 is equivalent to a $z$-score of 1; according to the normal curve, approximately 16% of the population exhibit scores in this range. Following the five-year
test point, the authors calculated a Global Deficit Score for each participant based on scores for all tests.

Findings indicated no impairment at any time for 67% of HCT survivors on digit symbol TMT-A or B. Between 22 and 38% of HCT survivors were impaired at pre-HCT or 80 days post on one or more tests, but not at one or five years post. Specifically, approximately 38% were impaired and recovered on the COWA, ≈20% on Digit Symbol, ≈38% on HVLT, 22% on HVLT-R, ≈30% on GPB-D, ≈32% on GPB-ND, ≈21% on TMT-A, and ≈22% on TMT-B. Neurocognitive dysfunction, as indicated by Global Deficit Scores, reportedly remained for 41.5% of survivors, compared with 19.7% of controls. More specifically, ≈15% of survivors remained impaired on COWA, ≈5% on Digit Symbol, ≈12% on HVLT, ≈15% on HVLT-R, ≈31% on GPB-D, ≈28% GPB-ND, ≈9% on TMT-A, and ≈8% on TMT-B. In the context of the normal curve, impairment in this sample was high overall. For those who remained impaired, motor speed as measured by the GPB-D was particularly high.

Overall, results suggested that survivors’ performance in information-processing speed and executive function continued to improve through five years post-transplant, although most HCT survivors had returned to pre-transplant functioning at one year post. Improvement on motor dexterity and verbal learning/retention did not improve beyond one year.

Interestingly, findings showed few differences between 5-year survivors and matched controls, with the exception of the survivors performing significantly worse on tests of motor speed and dexterity. Authors noted non-significant trends for better performance of matched-controls in verbal memory (measured via HVLT) and attention/visual scanning speed (measured via TMT-A).
Although study results may have been compromised by practice effects associated with traditional test measures, as well as attrition (only 66 completed the five-year analysis alongside controls), findings in general suggested continued recovery beyond one year post-transplant.

Differences between matched-control subjects underscored the subtle impairment associated with HCT.

**Delirium in HCT patients.** A handful of studies have also prospectively examined the incidence of delirium in HCT patients. Identification of delirium due to a medical condition is relevant to the study of neurocognitive change, given Diagnostic Statistical Manual of Mental Disorders (DSM-IV-TR) criteria that includes, among other features, a disturbance in consciousness and change in cognition, such as memory deficit, disorientation, and language disturbance (APA, 2000). However, the distinction between cognitive deficits and delirium must be recognized. Although memory deficits and language impairments are associated with both phenomena, delirium alone is fleeting and indicative of medical morbidity; thus, findings illustrating cognitive impairment in the context of delirium should be interpreted with caution. The following studies were included as they demonstrate the relationship between delirium and cognitive functioning in HCT patients and are thus considered both important and additive in their contribution to the HCT neuropsychological literature base.

In 2006, Beglinger et al. investigated delirium in HCT patients twice weekly with the DRS (Trzepacz et al., 1988), Memorial Delirium Assessment Scale (Breitbart et al., 1997) and brief measures of cognition during their inpatient stay. More than 40% of the sample exhibited delirium symptoms, predominantly occurring in the first two weeks following transplant. This finding is comparable to the prevalence of delirium in cancer (40%; Stiefel & Holland, 1991) and other medical patients (20-30%; Adamis, Treoloar, Martin, & Macdonald, 2006; Kaliswaart, de
Jonghe, & Bogaards, 2005). The presence of delirium in allogeneic transplant patients was found to be predictive of mortality. Other risks factors, however, such as age and cognitive status measures, did not predict death.

In a follow-up to this study, Beglinger and colleagues (2011) investigated both the occurrence of delirium in HCT patients as well as cognitive functioning before, during, and up to four weeks after HCT. The DRS and Memorial Delirium Assessment Scale were used to assess delirium; the Modified MMSE (3MS; Teng & Chui, 1987) was used to provide a measure of global cognitive functioning. A number of traditional neuropsychological tests were used in order to assess visual scanning, visuospatial/constructional abilities, and memory (TMT-A and TMT-B; Reitan & Wolfson, 1985), attention, language, visuospatial/constructional ability, and memory (Repeatable Battery for the Assessment of Neuropsychological Status [RBANS]; Randolph, Tierney, Mohr, & Chase, 1998), and estimated Full-Scale IQ (WASI; Psychological Corporation, 1999).

Assessments were undertaken twice weekly during patients’ inpatient stay in order to assess neuropsychological functioning and delirium status.

Findings indicated that within the first month of transplantation, 37% of survivors experienced an episode of delirium; for most of these patients, 84% experienced the episode within the first two weeks following transplantation. Findings indicated that deficits of those who experienced a delirium episode were most pronounced on measures of memory, learning, psychomotor speed, and attention. HCT patients who had not experienced a delirium episode performed between .5 and 1 SD below the mean on TMT-B and select RBANS tests (e.g. Coding, Semantic Fluency). Taken together, these findings provided additional support for the presence of neuropsychological deficit in HCT patients who have or have not experienced a
delirium episode. They likewise indicated that traditional neuropsychological testing may have utility for detecting delirium.

Fann, Alfano, Roth-Roemer, Katon, and Syrjala (2007) studied delirium and cognitive functioning in 90 HCT candidates, 81% of whom were scheduled for and subsequently underwent allogeneic transplantation. Participants completed assessments three times weekly, prior to transplantation and through 80 days post-transplant. Delirium was assessed via the DRS (Trzepacz et al., 1988) and neurobehavioral functioning via the Behavioral Dyscontrol Test (Grigsby, Kaye, & Robbins, 1992). Functioning was assessed in the following six domains: (a) cognitive flexibility (TMT-B [Reitan & Wolfson, 1985]), (b) visuomotor ability (Digit Symbol [Wechsler, 1997], (c) visual conceptual and visuomotor tracking ability (TMT-A [Reitan & Wolfson, 1985]), (d) immediate and delayed memory and learning ability (HVLT [Brandt, 1991]), (e) memory (Modified memory questionnaire [Sunderland, Harris, & Baddely, 1983]), (f) verbal fluency (COWA [Benton & Hamsher, 1983]), and (g) overall cognitive impairment (MMSE [Teng & Chui, 1987]).

Results revealed significantly worse executive functioning, attention, and processing speed at the concluding assessment, 80 days post-transplant, in those who had experienced a delirium episode. This study also illustrated that regardless of delirium status, at 30 days post-transplant, patients displayed decreased executive/frontal functioning. At 80 days post-transplant, patients who had not experienced an episode of delirium improved to above pre-transplantation levels, whereas those who had experienced delirium continued to have decreased executive/frontal neurobehavioral functioning. In light of these findings, in addition to minimizing medical morbidity, effective prevention or treatment of delirium and its underlying
etiologies and risk factors during HCT may be important in maximizing HCT recovery, including recovery of neuropsychological functioning.

Since this study was published in 2007, there have been two follow-up studies using the same data set. In 2010, Basinski and colleagues (2010) used study data to evaluate the impact of delirium on distress, health-related quality of life, and cognitive function six months and one year post-HCT. They found that patients who had experienced delirium within four weeks of HCT subjectively reported worse memory and executive functioning difficulty at 6 months and one year post-transplant. Fann and colleagues (2011) used data to evaluate pre- and post-transplant risk factors for delirium. These authors found that poorer pre-transplantation executive functioning was associated with more delirium severity. Taken together, these follow-up reports offer insight into the relationship between cognitive function and delirium; however, they are limited in that cognitive assessment was based on subjective cognitive function measures, as opposed to traditional neuropsychological tests.

Zaubler et al. (2010) examined the impact of delirium upon HCT patient decision-making in a prospective study of 19 HCT candidates, 90% of whom were scheduled for and subsequently underwent allogeneic transplantation. Participants were a subset of Fann and colleagues’ 2007 study. A variety of constructs were assessed, including delirium, decision-making capacity, and neurocognitive functioning. Measurement was completed prior to HCT through 80 days post. Results indicated that 36.8% of participants experienced a delirium episode, primarily starting between Day 7 and Day 16 post-transplant. Notably, early delirium was predictive of lower decision-making capacity at Day 30 post-transplant; results did not indicate decreased decision-making capacity at the initial, pre-transplant assessment.
Taken together, these findings provide limited information with regard to how delirium may impact neuropsychological function. Study results do confirm the presence of cognitive impairment in the absence of a delirium episode. However, the degree to which delirium may impact neuropsychological impairment remains unclear. At present, it appears that the experience of a delirium episode may predict later cognitive recovery post-transplant. Future work is needed to clarify whether the experience of delirium adds to HCT-associated neuropsychological impairment and, if so, to what extent.

**Other Factors Associated with Cognitive Decline**

There is a small literature identifying factors that may be associated with cognitive impairment in HCT patients. These include pre-existing factors such as older age (Booth-Jones et al., 2005; Padovan et al., 1998; Scherwath et al., 2012; Sostak et al., 2003; Syrjala et al., 2004) and lower premorbid intelligence (Booth-Jones et al., 2005; Scherwath et al., 2013; Sostak et al., 2003).

Prior treatment regimens have also been linked to the cognitive impairment exhibited in HCT candidates and survivors. Syrjala and colleagues (2004) identified a higher risk for cognitive deficits for those with a prior history of chemotherapy than for those without. Furthermore, a number of researchers have shown a history of cranial radiation to be associated with impairment (Andrykowski et al., 1992; Booth-Jones et al., 2005; Sostak et al., 2003).

**Risk associated with HCT pre-conditioning regimens.** Various HCT pre-conditioning regimens have also been shown to be associated with cognitive impairment. Intrathecal chemotherapy and total body irradiation, particularly when combined with central nervous system disease, were shown to be independent predictors of poor test performance by Andrykowski and colleagues (1992).
Others have not found a correlation between treatment regimen and cognitive impairment. Ahles and colleagues (2001) found no differences between the cognitive functioning of those who had a history of radiation and/or intrathecal therapy and those without this treatment history. The authors noted that this may have been due to the small number of study participants who had received these prior treatments. Sostak and colleagues (1998) likewise found no association between cognitive impairment and total body irradiation.

Schultz-Kindermann et al. (2007) also found no significant differences in cognitive functioning between groups receiving standard or reduced-intensity pre-HCT conditioning regimens. These findings should be interpreted with caution, however, given the small sizes of comparison groups within the sample.

**GVHD-related cognitive deficits.** In general, literature indicates increased cognitive impairment in allogeneic patients in comparison with those who have undergone and autologous transplant. This may be due to the impact of acute or chronic GVHD.

Jacobs, Small, and Booth-Jones (2007a) identified increased cognitive deficits in allogeneic HCT patients with GVHD as compared to those without. Their sample included post-transplant patients ($N = 388$), 21% of whom underwent allogeneic transplantation. GVHD subsequently developed in 13% of the sample, although study researchers did not report whether participants’ GVHD diagnoses were acute or chronic. Participants completed assessments of (a) memory (CVLT [Delis et al., 1996], Logical memory and Visual Reproduction [Wechsler, 1997]), (b) executive functioning (Digit symbol [Wechsler, 1981], TMT-A, TMT-B [Reitan & Wolfson, 1985], COWA [Benton & Hamsher, 1983], Stroop [Stroop, 1935]), (c) attention (CPT-II [Conners, 2000]), and (d) motor functioning (GPB [Kløve, 1964]). Those with GVHD were
found to perform significantly worse \((p < .05)\) on memory and psychomotor speed; they also demonstrated an overall lower total neuropsychological performance than those without GVHD.

This increased risk for cognitive impairment may also be due, in part, to immunosuppressant therapy, used in the treatment of GVHD. Syrjala and colleagues (2004) found that patients who had received immunosuppressants, including cyclosporine and others (e.g., tacrolimus or mycophenolate mofetil), demonstrated increased motor impairment as compared to those without exposure to treatment with immunosuppressants.

These results are supported by other studies that have linked cognitive impairment to chronic GVHD in particular (Booth-Jones et al., 2005; Sostak et al, 2003) and the long-term immunosuppressant therapy used in its treatment. Padovan et al. (1998) found an association between long-term corticosteroid therapy and pathological MRI results, white matter lesions, and neuropsychological impairment.

**Neuropsychological Domains Affected**

Taken together, the literature suggests that a number of HCT-associated risk factors may contribute to neuropsychological impairment. Collectively, the impact is broad. Neuropsychological impairment demonstrated in HCT candidates before transplantation has been shown to occur primarily within memory, psychomotor functioning, attention, and executive function (Andrykowski et al., 1992; Beglinger et al., 2006; Booth-Jones et al., 2005; Harder et al., 2002, 2005, 2006; Jacobs, 2007; Meyers et al., 1994; Schultz-Kindermann et al., 2007; Syrjala, 2004), though deficits in verbal fluency and visuomotor skills have also been noted (Booth-Jones, 2005; Harder, 2002, 2005, 2006; Syrjala, 2004).

Some literature suggests that this impairment increases through approximately 3 months post-transplant (Beglinger et al., 2006; Syrjala et al., 2004) before improving, although other
findings suggest that impairment rises through 8 month post-transplant (Meyers, 1994). In general, cognitive deficits in HCT patients have been shown to return to pre-transplant levels at one year post-transplant (Jacobs et al., 2007; Syrjala et al., 2004), although other work has demonstrated a mere stabilization of cognitive decline (Harder et al., 2005).

Little is known about long-term neuropsychological functioning. Although Harder et al. (2002) did demonstrate impairment in HCT survivors long after transplant (22-82 months post-transplant), it is unclear how much impairment may have changed across the transplant trajectory. Syrjala and colleagues (2011) prospectively evaluated HCT survivors through five years post-transplant and showed cognitive improvement at five years post-transplant. However, use of traditional tests may have confounded results of this study, given the possibility of practice effects. Future studies are warranted, particularly using measures with minimal practice effects, in order to expand on existing literature and clarify the degree to which cognitive deficits persist for HCT survivors.

**Alternatives to Traditional Neuropsychological Measurement**

It is possible that challenges associated with traditional neuropsychological tests have had an impact on the dearth of literature addressing long-term functioning. Given that traditional tests must be administered by a neuropsychologist or highly trained psychometrician, costs associated with longitudinal study designs using repeated assessment are high. Further, many traditional neuropsychological tests designed to detect subtle cognitive changes are susceptible to practice effects, rendering the results of repeated assessment invalid when repeated administrations are conducted less than six months apart. As such, many potential investigators may be discouraged from research efforts in this domain.
In order to address some of the difficulties associated with traditional measurement in HCT, some researchers have recently turned to alternative, cost- and time-efficient neuropsychological measures.

**Subjective cognitive self-evaluation.** Measurement of subjective cognitive functioning has been explored as one such option. The Functional Assessment of Cancer Therapy Cognitive Measure (FACT-Cog; Wagner, Lai, Cella, Sweet & Forrestal, 2004), the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC-QLQ; Aaronson et al., 1993) and Multiple Abilities Questionnaire (MAQ; Seidenberg et al., 1994) are tools that have recently been used to assess patient-reported cognitive deficits in HCT patients. Other studies have made use of interviewing techniques to quantify patient subjective cognitive function (Harder et al., 2002, 2005, 2006).

Although the FACT-Cog (Wagner et al., 2004) was designed specifically to capture the cognitive experience of cancer patients, Jacobs and colleagues (2006) found FACT-Cog subscales to be largely unrelated to objective measures, with the exception of one subscale: “Other People Noticed Deficits.” Scores on this subscale were significantly correlated with actual cognitive performance, suggesting that this dimension of perceived cognitive functioning (i.e., the subjective experience of others pointing out potential deficits) could be used as a reliable measure of brief neuropsychological assessment.

The MAQ (Seidenberg et al., 1994) likewise appears to lack the overall sensitivity necessary to detect subtle cognitive changes associated with HCT. Booth-Jones and colleagues (2005) found that participants’ complaints of remote memory, attention/concentration, and language captured via MAQ measurement were not significantly associated with total neuropsychological performance. MAQ-measured visual-spatial perception, visual-spatial
memory, and remote memory, however, were significantly correlated with other domains subjectively measured. MAQ-measured visual memory performance was likewise significantly correlated with objective measurement of objective visual-spatial perception measurement. Notably, participants’ complaints were related to depressive symptomology and fatigue ratings.

Similar problems have emerged through use of the EORTC (Aaronson et al., 1993) to assess neuropsychological function. Although this instrument includes a cognitive functioning scale, Harder and colleagues (2002) found subscales measuring global health, education level, and fatigue to be the strongest predictors of cognitive test impairment, rather than the cognitive function scales.

Interviewing techniques for assessing cognitive impairment likewise appear to provide limited information with regard to objective cognitive functioning. Using a clinical interview to collect information from participants, Harder and colleagues (2002) found no correlation between reported experiences and cognitive function measured via traditional techniques. This mirrors other work suggesting little (Poppelreuter et al., 2007) to no (Harder et al., 2005, Schulz-Kindermann et al., 2007) correlation between neuropsychological testing and patient self-evaluation.

Overall, it remains unclear whether this disconnect is due to patient reporting errors associated with diminished awareness or whether traditional tests lack the needed sensitivity to capture subtle changes in cognition.

**Computerized cognitive testing.** Computerized cognitive testing also presents as an alternative to traditional neuropsychological testing. A number of advantages are associated with computerized neuropsychological assessment batteries, including increased cost-effectiveness with regard to data scoring and analysis, capacity for complex stimuli generation, increased
Computerized testing is not without its own set of unique problems. The presence of visuo-perceptual demands in computerized tests may cause difficulty for some. Likewise, a person’s overall familiarity with computers may influence performance (Kapur, 1988). Further, if administration and result interpretation is done in the absence of licensed clinical neuropsychologist, interpretation errors may also present as a threat.

Recently, however, the use of computerized, web-based memory testing for preliminary screening and repeated testing over time to detect change has been recognized as a promising alternative, if done under the supervision of a neuropsychologist (Crook et al., 2009). MicroCog (Powell et al., 1996), a computerized assessment tool, was recently piloted in elderly cancer patients with at least minor visual impairment, the majority of whom had never used a computer. Results suggested that computer-based testing may be a feasible tool for detecting neuropsychological comorbidity in older cancer patients (Extermann, 2005). Likewise, CogState, a computerized neuropsychological assessment battery, designed for repeated assessment with minimal practice effects, has been validated in a variety of settings and has been suggested as a tool for use in evaluating changes in chemotherapy-associated cognitive functioning across large, multicenter clinical trials (Tannock et al., 2004).

**CogState** (Collie, Maruff, & Darby, 2003). “CogState” refers to a battery of computerized neuropsychological measures of visuomotor function, psychomotor/processing speed, visual attention, attention/working memory, verbal learning and memory, executive function, and social cognition. Test developers recognized several benefits associated with assessing performance changes over very short test-retest intervals, as these circumstances may
clarify the extent to which a testing subject has acclimated to a test and how much practice may affect performance over time (Falleti, Maruff, Collie, and Darby, 2006).

One study investigating practice effects associated with CogState examined healthy individuals’ performance at very brief test-retest intervals (e.g. 10-minutes and one week). Results indicated that performance improved from the first to second assessment on the test battery; following the second assessment, performances stabilized. Further, test performance over the first four repeated measurements did not worsen, a condition purported to arise if individuals had become fatigued or lost motivation secondary to the repeated testing. Thus, the authors concluded that beyond the initial two assessments, significant increases in scores could be seen to reflect actual cognitive improvement rather than improvement due to practice. They recommended that initial practice effects be controlled for through the use of dual baseline assessment and statistical procedures (Falleti et al., 2006). Similar results were also reported in an earlier study by Collie, Maruff, Darby, and McStephen (2003), who evaluated cognitive impairment associated with 24 hours of sustained wakefulness and blood alcohol concentration of .05% in healthy subjects, using repeated CogState test sessions. This study also found that practice effects were mainly observed between the first and second test administrations and statistically non-significant improvements occurring thereafter.

Work investigating the utility of CogState Brief Battery specifically has shown temporal stability in at least some of the tasks. When used in a pharmacological study, the Detect, Identify, and Learn tasks specifically demonstrated sensitivity to sedative effects of the agent, midazolam, with no practice effects observed. Practice effects were, however, observed in participant performance on the Groton Maze Learning Task, considered to be more cognitively demanding (Collie et al., 2007). Another investigation by Lim and colleagues (2013) demonstrated high test-
retest reliability in various CogState tasks over a three-month time period. These authors investigated the utility of six CogState tasks administered to 105 healthy older adults, 48 adults with mild cognitive impairment, and 42 adults with Alzheimer’s disease. Tasks included Detect, a measure of processing/psychomotor speed; Identify, a measure of visual attention; One-Back, a measure of work memory; Continuous Paired Associative Learning (CPAL) and Learn, measures of visual learning and memory; and the International Shopping List (ISL) task, a measure of verbal memory and learning. Results indicated that CogState tasks possessed the sensitivity to detect varying degrees of cognitive impairment, such that mild cognitive impairment and dementia patients performed more poorly than healthy older adults. Tasks also exhibited stability over three monthly test points.

Finally, test-retest reliability of various CogState tasks (i.e., Detect, Identify, One-Back, Learn, and CPAL) was also evidenced by Hammers and colleagues (2011), who investigated the battery’s utility in dementia patients.

CogState has been validated for use with clinically sensitive populations, including patients with HIV-1 infection (Cysique et al., 2006) and dementia (Hammers et al. 2011). Furthermore, it has shown sensitivity in detecting subtle cognitive deficits related to fatigue, intoxication, and stress (Falleti et al., 2003).

One specific battery of CogState tasks was validated in AIDS dementia, schizophrenia, and traumatic brain injury groups (Maruff et al., 2009) showing particular promise. Known as the CogState Brief Battery, the task set includes measures of processing/psychomotor speed (Detect), visual attention (Identify), working memory (One-Back), and visual learning and memory (Learn). Validation included assessments of construct and criterion validity (2009). Construct validity was evaluated by examining healthy controls’ performance on CogState and
traditional neuropsychological tasks; as such, in addition to CogState tasks, participants completed traditional neuropsychological tests, which included GPB, TMT, SDMT, BVMT, RCFT-R, and Span (see Table 1 for a list of tests and abbreviations). Pearson product-moment correlations between tasks were computed and transformed into effect sizes. Effect sizes were plotted with their 95% confidence intervals in order to determine the strongest associations between tasks. Results indicated that Detect was most strongly associated with measures requiring simple visual attention or psychomotor functions (GPB and TMT-A). Participant performance on Identify was most strongly associated with measures requiring visual processing and divided attention (TMT and SDMT). One-Back was most strongly associated with tasks requiring visual scanning and working memory (SDMT and Span), and Learn most strongly with tasks requiring visual learning and memory (BVMT and RCFT). The One-Back task was most strongly associated with measures requiring visual scanning and working memory (SDMT and Span). Therefore, it was determined that the battery possessed acceptable construct validity.

Next, criterion validity was evaluated through establishing the magnitude of impairment of criterion samples’ impairment as compared with healthy controls. Traumatic brain injury, schizophrenia, and AIDS dementia groups were chosen as criterion samples, in that these groups are characterized by cognitive impairment in attention, processing speed, memory, and executive function. Furthermore, although impairment in these domains is generally disruptive to patients’ living, it does not limit neuropsychological measurement procedures. Thus, investigators believed the groups to be ideal for evaluating CogState criterion validity. It was hypothesized that CogState Brief Battery tasks would possess sufficient criterion validity to detect differences between criterion and control groups as well as the capability to accurately predict [criterion vs. control] group membership. In order to test these hypotheses, a series of independent t-tests were
first performed to assess differences between criterion and control group performances on each CogState Brief Battery task. Difference scores were next transformed into effect sizes in order to assess the magnitude of differences. Together with associated 95% confidence intervals (CIs), the non-overlap (%NOL) for each task was computed. The %NOL statistic reflects the extent to which data distributions of clinical and control groups do not overlap (i.e., the proportion of clinical group scores not shared by the control group). According to Zakzanis (2009), a %NOL statistic of 93% is considered to be a good clinical marker. Results of independent $t$–tests indicated that differences for all group tasks were statistically significant between control and clinical groups except for the Detect task within the TBI group. However, further evaluation with the %NOL statistic revealed less than ideal predictive capability.

Within the three criterion groups, %NOL was on average, between 50-60%. Overall, %NOL for the Learn task was highest, ranging from 52-78%. For the Detect task, %NOL ranged from 41-62%; %NOL on the Identify task ranged from 53-62%, and on the One-Back, 55-60%. Investigators posited that predictive validity was lower than anticipated due to the subtlety of cognitive impairment associated criterion samples. In turn, they suggested that comparisons using a criterion sample with greater impairment might yield higher %NOL. It was concluded that although, independently, CogState Brief Battery tasks failed to possess criterion validity needed to predict group membership, they nevertheless demonstrate sensitivity to cognitive change.

A recent investigation of CogState tasks in a Japanese schizophrenic sample provided additional support for the utility of this battery. Yoshida and colleagues (2011) investigated the criterion and construct validity of Detect, Identify, One-back, Learn, CPAL, GML, ISL, and the SECT, a social cognition task, collectively known as the CogState Schizophrenia Battery. Forty
patients with schizophrenia underwent testing with the Japanese version of this battery as well as the Brief Assessment of Cognition in Schizophrenia (BACS; Keefe et al., 2004). The means of CogState scores were added and averaged in order to compute a composite neuropsychological score, which was compared with a BACS composite score. Results exhibited a strong relationship between BACS and CogState Schizophrenia Battery scores ($r = .71$, $p < .001$) and, thus, acceptable construct validity. Magnitudes of impairment (i.e., $z$-scores) for schizophrenic patients were examined against scores of matched healthy controls participants. Results revealed a distinct pattern of impairment, which the authors reported as evidence of criterion validity. A factor analysis performed on this CogState battery suggested that three scores of cognitive performance could be derived from the scores, including (a) memory (including CPAL, Learn, ISL, and GML tasks), (b) speed of performance (including Detect and Identify tasks), and (c) social cognition, from the SECT task. Although further criterion validity work is warranted for the use of CogState tasks used in this battery, the study overall yielded additional support for the utility of CogState.

These criterion validity findings are consistent with Hammers and colleagues (2011), who also found CogState to effectively discriminate between dementia patients ($n = 91$) and healthy controls ($n = 20$); discriminative ability was less strong in other areas. This study aimed to investigate the short-term stability and reliability of various CogState tasks, including Detect, Identify, One-Back, and Learn, as well as two additional card tasks, over two hour-apart test sessions. The tasks were administered to 23 healthy control subjects, 20 with mild cognitive impairment, 20 with Alzheimer’s disease, 10 with dementia with lewy bodies, and 9 with frontal lobe dementia. Analyses of task variance between across groups demonstrated that although CogState tasks could effectively discriminate between dementia and healthy control subjects,
tasks were unable to discern dementia patients from each other (i.e., dementia with lewy bodies from frontal lobe dementia). Tasks were furthermore unable to discern mild cognitive impairment subjects from healthy control subjects. The tasks did perform well with regard to test-retest reliability and practice effects were minimal over the course of the two test administrations, but study results overall suggested less than optimal task discriminability.

CogState validity in general is still in question, however. Other results by Overton et al. (2011) suggest that the battery has less than optimal construct validity. These authors found lower than expected relationships between CogState tasks and traditional neuropsychological assessment measures. Findings indicated strongest relationships between speeded CogState measures and traditional tests. More specifically, (a) Detect, described by authors as a measure of simple reaction time for assessing psychomotor function and speed of processing domains and (b) Identify, described by authors as a measure of simple reaction time, were most strongly correlated with global impairment scores calculated from means of traditional tests. Correlations between CogState tasks and traditional tests were in general, however, weak. Notably, the authors reported a weak relationship between TMT-B and CogState tasks, in contrast to findings by Cysique and colleagues (2006).

Overall, these investigations into the reliability and validity of CogState suggest that various tasks, particularly those within the CogState Brief Battery, appear to have utility in repeated assessment. Temporal stability appears to be a strength of CogState tasks, although further evaluation surrounding construct and criterion validity is warranted.

Conclusion

In conclusion, the present literature base demonstrates a pattern of neuropsychological change across the HCT trajectory. Impairment has been noted in HCT candidates prior to
transplantation and appears to persist before improving approximately 100 days following HCT. Impairment has been shown in both those with and without a history of delirium and in a number of cognitive domains across the transplant trajectory, including executive functioning, memory, attention, verbal learning and psychomotor speed. Although some literature has illustrated an association between cognitive impairment and pre-transplant conditioning regimens, many studies are limited by methodological issues including small sample sizes and attrition due to medical morbidity and mortality. Collectively, over 60 measures of neuropsychological function have been used to measure cognitive function in HCT patients and therefore, results are difficult to comparatively quantify. Further complicating the literature, studies have employed a multitude of variety of criteria for classifying impairment, and many investigators have noted limitations associated with the potential for practice effects. These effects are important to consider in tracking cognitive recovery because they have the potential to confound study results, as neuropsychological improvements demonstrated may be due to practice rather than actual neuropsychological recovery.

Aims of investigation

Rationale

Overall, few studies have evaluated cognitive impairment across the HCT transplant trajectory. Although cross-sectional work has evidenced cognitive deficits at key intervals across the HCT trajectory, the use of multiple assessment instruments and variable impairment criteria has made results difficult to quantify. Prospective studies, which capture cognitive change over time, have used repeated measurements. However, given the possibility of practice effects associated with many traditional tests, findings of these prospective studies are potentially confounded, such that practice effects would mask cognitive impairment occurring at later
assessment times. Consequently, further prospective work is necessary in order to clarify the clinical picture, particularly with reliable measurement that is appropriate for repeated use.

Given its capability for repeated assessment with minimal practice effects after the second administration, we hypothesized that use of the CogState Brief Battery could be helpful in tracking cognitive changes associated with HCT. Conveniently, tasks included within the CogState Brief Battery (e.g., Detect, Identify, One-Back, and Learn) are representative of key cognitive domains that are frequently impaired in HCT patients (processing/psychomotor speed, visual attention, working memory, and visual learning and memory, respectively). Thus, the present study proposed to evaluate validity of the CogState Brief Battery for use in HCT patients. However, given that HCT patients and survivors also experience impairment in executive function and verbal fluency, the CogState Brief Battery was expanded to include supplemental tasks from the full CogState test battery to measure these domains. Specifically, the following tasks were added: the Groton Maze Learning (GML) task, a measure of executive function; the International Shopping List (ISL) task, a measure of verbal learning; and the Continuous Paired Associative Learning (CPAL) task, a visual learning and memory task which discriminated well in a breast cancer population (B. Giordani, personal communication, 3/19/2010). Going forward, this expanded battery will be referred to as the Adapted CogState Brief Battery.

The research method of the present study replicated Maruff and colleagues’ 2009 study, which assessed CogState Brief Battery construct and criterion validity. The present study was undertaken so that if found to be a valid measurement tool for use within the HCT population, the Adapted CogState Brief Battery could be used to study longitudinal neuropsychological change in HCT candidates. Further, it could help clarify the degree to which HCT survivors experience neuropsychological changes over time.
Research Questions

As previously noted, the present study sought to investigate construct and criterion validity of the CogState tasks (i.e., Adapred CogState Brief Battery).

Construct validation is important as it provides a measurement of the relationship of new test variables to traditionally used instruments. As per Shadish, Cook, and Campbell (2002), construct validity is the degree to which inferences are warranted from the observed persons, settings, and cause and effect operations included in a study to the constructs that these instances might represent. Thus in order to determine construct validity, each CogState task (e.g., Detect, Identify, One-Back, Learn, CPAL, GML, ISL) was examined against traditional neuropsychological assessment measures (e.g., GPB, TMT, SDMT, Span, RCFT-R, HVLT) through calculating a series of Pearson correlation coefficients. Criterion validity is likewise important, referring to the extent to which results on one measure are associated with results from another measure considered to be the criterion variable (Streiner & Norman, 1995). As defined by Pedhazur and Schmelkin (1991), a criterion refers to a variable that can be explained or predicted by information from another variable. Criterion validity of the Adapted CogState Brief Battery was evaluated by examining the degree to which CogState tasks could discern HCT survivors from healthy spousal partners through calculation of the %NOL and a series of exploratory paired t-tests.

Hypotheses

Two primary sets of hypotheses were proposed to address research questions.

1. **Construct validation hypotheses.** First, construct validity of the Adapted CogState Brief Battery was examined to determine the degree of association between CogState and traditional neuropsychological tasks. It was hypothesized that Adapted CogState Brief
Battery tasks would correlate significantly with traditional tests purported to measure similar neuropsychological domains. See Table 2.

Table 2

<table>
<thead>
<tr>
<th>Domain</th>
<th>CogState</th>
<th>Traditional</th>
<th>Expected correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Task</td>
<td>Score direction</td>
<td>Task</td>
</tr>
<tr>
<td>a. Processing/ Psychomotor Speed</td>
<td>Detect</td>
<td>Lower</td>
<td>GPB-D</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GPB-ND</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TMT-A</td>
</tr>
<tr>
<td>b. Visual Attention</td>
<td>Identify</td>
<td>Lower</td>
<td>TMT-A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TMT-B</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SDMT</td>
</tr>
<tr>
<td>c. Working memory</td>
<td>One-Back</td>
<td>Higher</td>
<td>SDMT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Span</td>
</tr>
<tr>
<td>d. Visual Learning and Memory</td>
<td>Learn</td>
<td>Higher</td>
<td>RCFT-R</td>
</tr>
<tr>
<td></td>
<td>CPAL</td>
<td>Lower</td>
<td></td>
</tr>
<tr>
<td>e. Executive Function</td>
<td>GML</td>
<td>Lower</td>
<td>TMT-B</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SDMT</td>
</tr>
<tr>
<td>f. Verbal Learning and Memory</td>
<td>ISL</td>
<td>Higher</td>
<td>HVLT</td>
</tr>
</tbody>
</table>

Note: Direction of good performance for all traditional tests were corrected in standardization, so higher is better. This was not the case for CogState tasks; direction of good performance is listed above alongside the expected correlation directions.

a. **Processing/psychomotor speed.** Performance on the CogState Detect task, a measure of processing/psychomotor speed, was hypothesized to demonstrate the strongest associations with tasks requiring simple visual attention, or psychomotor functions: GPB-D, GPB-ND, and TMT-A.
b. Visual attention. Performance on the Identify task, a measure of visual attention, was hypothesized to demonstrate the strongest associations with tasks requiring visual processing and divided attention: TMT-A, TMT-B, and SDMT.

c. Working memory. Performance on the One-Back task, a measure of working memory, was hypothesized to demonstrate the strongest associations with tasks requiring visual scanning and working memory: SDMT and Span.

d. Visual learning and memory. Performance on the Learn task, a measure of continuous visual recognition learning, was hypothesized to demonstrate the strongest associations with tasks requiring visual learning and memory: RCFT.

e. Executive functioning. Performance on the GML, a measure of executive functioning, was hypothesized to demonstrate the strongest associations with tasks requiring visual processing and mental flexibility: SDMT and TMT-B.

f. Verbal learning and memory. Performance on the ISL, a measure of verbal learning and memory, was hypothesized to demonstrate the strongest associations with tasks requiring verbal learning and memory: HVLT, HVLT-R, HVLT-RET

2. Criterion validation hypotheses. A second series of hypotheses was formulated with regard to the Adapted CogState Brief Battery criterion validity. Healthy control participant performance on CogState tasks was expected to differ significantly from HCT survivor performance. Based on findings of Maruff and colleagues’ prior study (2009), as measured by the %NOL statistic, tasks were expected to yield average %NOLs of 50-60%. Criterion sample (HCT survivor) performance was expected to differ significantly from control performance (spousal partner) on CogState tasks. Overall, CogState tasks were hypothesized to yield $\geq$ %NOL for tasks previously evaluated by Maruff et al. (2009).
a. **Processing/psychomotor speed.** Detect was hypothesized to yield a %NOL of 41-62% or greater.

b. **Visual attention.** Identify was hypothesized to yield a %NOL of 53-61% or greater.

c. **Working memory.** One-Back task was hypothesized to yield a %NOL of 55-60% or greater.

d. **Visual learning and memory.** Learn was hypothesized to yield a %NOL of 52-78% or greater.

e. **Supplemental tasks: visual learning and memory, executive function, verbal learning and memory.** Supplemental CogState tasks outside of the Brief Battery (e.g., CPAL, GMLT and ISLT) were hypothesized to yield a %NOLs of 41-78%.

Exploratory analyses were also used in order to evaluate discriminative ability of CogState: a series of paired t-tests were computed in order to evaluate the ability of CogState tasks to discriminate between survivors and spousal partners. Significant differences between paired groups would be seen as indicative of strong discriminative ability. It was hypothesized that CogState tasks would show strong discriminative ability, evidenced by statistically significant differences between pairs ($p < .05$).

**Methods**

**Participants**

Given the increased neuropsychological deficits associated with allogeneic transplantation (Padovan et al., 1998; Booth-Jones et al., 2005; Jacobs et al., 2007a), it was originally proposed that only allogeneic survivors within three years of transplant be included within the study; it was anticipated that, for the purposes of evaluating criterion validity, allogeneic survivors would demonstrate enough impairment to detect notable differences from
partner control subjects. However, recruitment of subjects meeting these criteria became challenging, so inclusion criteria were expanded to include both autologous and allogeneic survivors within five years of transplant.

Participants were recruited through local survivor support programs such as Gilda’s Club, cancer support groups (including on-site, phone, internet-based groups), and workshops jointly hosted by the Eastern Michigan University Psychology Department and Leukemia and Lymphoma Society, featuring presenters from the Eastern Michigan University Behavioral Medicine Research Group. Within each of these venues, the principal investigator presented information about the study via spoken and written information (see Appendix K). Program attendees were invited to participate in the study if they had undergone an HCT within the last five years, were ≥ 18 years of age, were able to read and speak English, and had a spouse or domestic partner who also agreed to participate in the study. They were also required to live within 500 miles of Detroit in order to participate. Study candidates were excluded from enrollment if they did not meet aforementioned eligibility criteria (see Appendix A) or if they reported any of the following, which could additively affect neuropsychological test performance: an existing neurological disorder; prior head injury, past or current substance abuse and/or dependence; history of stroke, epilepsy, or other CNS pathology requiring radiation or surgery; or history of brain tumor(s).

Contact information was collected from persons who expressed interest in participating, and the principal investigator later contacted them in order to re-review all study criteria, provide information about participation, and schedule a time for both the HCT survivor and spouse to complete study procedures in their home, local library, or at the Eastern Michigan University Psychology Clinic.
A power table was consulted to determine the number of participants needed. Using a large effect size of .4, two predictor variables (i.e., group, premorbid IQ), and following conventions for probability and desired power level (.05 and .8, respectively), the analysis revealed a necessary minimum of 26 participants (Cohen, 1992). Given that the present study replicated prior validation efforts by Maruff and colleagues (2009), a sample of 20 dyads was expected to provide sufficient power for detecting group differences, provided that group variability be reduced through co-varying out difference due to premorbid intelligence.

A total of 20 dyads \((N = 40)\) participated in this study. Of the 20 participating dyads, 19 were married; one was a domestic partnership. Ninety-five percent were White and all relationships were heterosexual. Participants were a mean of 53 years of age and had completed a mean of 15 years of education. On average, participants achieved a mean pre-morbid IQ score of 108.34, as estimated by the NAART. See Table 3 for a summary of sample demographic characteristics.
Table 3

Summary of demographic characteristics

<table>
<thead>
<tr>
<th>Demographics</th>
<th>HCT survivors</th>
<th>Spousal partners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Mean age</td>
<td>52.65 (7.25)</td>
<td>53.3 (6.63)</td>
</tr>
<tr>
<td>Gender (% males)</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>Mean years of education</td>
<td>15.10 (1.917)</td>
<td>14.35 (2.03)</td>
</tr>
<tr>
<td>% sample self-identifying as White</td>
<td>95</td>
<td>95</td>
</tr>
<tr>
<td>% sample self-identifying as Native American</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>% sample self-identifying as Middle Eastern</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Estimated pre-morbid IQ(^a)</td>
<td>108.65 (7.12)</td>
<td>108.03 (7.96)</td>
</tr>
</tbody>
</table>

Employment status at participation

<table>
<thead>
<tr>
<th>Employment status</th>
<th>n (%)</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employed</td>
<td>6 (30)</td>
<td>12 (20)</td>
</tr>
<tr>
<td>Full-time</td>
<td>4 (20)</td>
<td>10 (5)</td>
</tr>
<tr>
<td>Part-time</td>
<td>2 (10)</td>
<td>2 (15)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>3 (15)</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Due to transplant-related issues</td>
<td>3 (15)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Full-time student</td>
<td>0 (0)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Retired</td>
<td>4 (20)</td>
<td>3 (15)</td>
</tr>
<tr>
<td>Due to transplant-related issues</td>
<td>2 (10)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>For other reasons</td>
<td>2 (10)</td>
<td>3 (15)</td>
</tr>
<tr>
<td>On disability</td>
<td>7 (35)</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Due to transplant-related issues</td>
<td>5 (25)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>For other reasons</td>
<td>1 (5)</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Medical leave due to transplant</td>
<td>1 (5)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

Self-reported socioeconomic status

<table>
<thead>
<tr>
<th>Socioeconomic status</th>
<th>n (%)</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>We have barely enough to get by</td>
<td>2 (10)</td>
<td>3 (15)</td>
</tr>
<tr>
<td>We have enough to get by, but no more</td>
<td>5 (25)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>We are solidly middle class</td>
<td>9 (45)</td>
<td>11 (55)</td>
</tr>
<tr>
<td>We have plenty of extras</td>
<td>2 (10)</td>
<td>4 (20)</td>
</tr>
<tr>
<td>We have plenty of luxuries</td>
<td>1 (5)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Don't know/unsure/prefer not to say</td>
<td>1 (5)</td>
<td>1 (5)</td>
</tr>
</tbody>
</table>

Annual household income

<table>
<thead>
<tr>
<th>Income range</th>
<th>n (%)</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 150,000</td>
<td>1 (5)</td>
<td>3 (15)</td>
</tr>
<tr>
<td>100,000 - 149,000</td>
<td>5 (25)</td>
<td>3 (15)</td>
</tr>
<tr>
<td>75,000 - 99,000</td>
<td>3 (15)</td>
<td>4 (20)</td>
</tr>
<tr>
<td>50,000 - 74,000</td>
<td>4 (20)</td>
<td>3 (15)</td>
</tr>
<tr>
<td>25,000 - 49,000</td>
<td>3 (15)</td>
<td>4 (20)</td>
</tr>
<tr>
<td>10,000 - 24,000</td>
<td>2 (10)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Don't know, or prefer not to say</td>
<td>2 (10)</td>
<td>2 (10)</td>
</tr>
</tbody>
</table>

\(^a\)Estimated pre-morbid IQ derived from National Adult Reading Test
Survivors participating in the study were on average two years and six weeks post-transplant, ranging from 10 weeks to five years, three months post-transplant the time of participation. The majority of survivor participants were diagnosed with acute myelogenous leukemia (AML; 35%), followed by multiple myeloma (25%). More than half of our sample received an allogeneic transplant. High dose chemotherapy was received by 95% of survivors. Sixty percent of our sample suffered graft-versus host disease (GVHD) at some point during their transplant trajectory. See Table 4 for a summary of survivor diagnoses, treatment regimen, GVHD prevalence, and GVHD treatment received.
Table 4

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute Myelogenous Leukemia</td>
<td>7 (35)</td>
</tr>
<tr>
<td>Chronic Lymphocytic Leukemia</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Chronic Myelogenous Leukemia</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Multiple Myeloma</td>
<td>5 (25)</td>
</tr>
<tr>
<td>Myelodysplastic Syndrome</td>
<td>4 (20)</td>
</tr>
<tr>
<td>Non-Hodgkins Lymphoma</td>
<td>2 (10)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Transplant type</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autologous</td>
<td>7 (35)</td>
</tr>
<tr>
<td>Allogeneic</td>
<td>13 (65)</td>
</tr>
<tr>
<td>Unrelated, matched donor</td>
<td>6 (30)</td>
</tr>
<tr>
<td>Related, matched donor</td>
<td>6 (30)</td>
</tr>
<tr>
<td>Cord blood donor</td>
<td>1 (5)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Preconditioning regimen</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total body irradiation</td>
<td>6 (30)</td>
</tr>
<tr>
<td>High dose chemotherapy</td>
<td>19 (95)</td>
</tr>
<tr>
<td>Intrathecal chemotherapy</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Systemic chemotherapy</td>
<td>7 (35)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GVHD prevalence</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No GVHD</td>
<td>8 (40)</td>
</tr>
<tr>
<td>GVHD</td>
<td>12 (60)</td>
</tr>
<tr>
<td>Acute</td>
<td>3 (15)</td>
</tr>
<tr>
<td>Acute which developed into chronic GVHD</td>
<td>3 (15)</td>
</tr>
<tr>
<td>Late-onset chronic GVHD</td>
<td>5 (25)</td>
</tr>
<tr>
<td>Other difficulties treated with</td>
<td>1 (5)</td>
</tr>
<tr>
<td>immunosuppressants</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>GVHD Medications</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyclosporine</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Tacrolimus</td>
<td>7 (35)</td>
</tr>
<tr>
<td>Mycophenolate mofetil</td>
<td>3 (15)</td>
</tr>
<tr>
<td>Glucocorticoids</td>
<td>5 (25)</td>
</tr>
<tr>
<td>Sirolimus</td>
<td>3 (15)</td>
</tr>
<tr>
<td>Methotrexate</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Other medications</td>
<td>3 (15)</td>
</tr>
<tr>
<td>Gleevac</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Steroid cream</td>
<td>2 (10)</td>
</tr>
</tbody>
</table>
Procedures

Prior to participation, the informed consent form was reviewed. It was based upon a script (see Appendix B) and consent form (see Appendices C-D), which provided a review of study intent, inclusion and exclusion criteria, procedures involved, and confidentiality stipulations.

After consenting to participate, participants first completed a series of brief questionnaires (described below), including the Patient Health Questionnaire (PHQ-2; Kroenke, Spitzer, & Williams, 2003), Generalized Anxiety Disorder-2 (GAD-2; Kroenke, Spitzer, Williams, Monahan, & Lowe, 2007), Sleep Problems Questionnaire (SPQ; Jenkins, Stanton, Niemcryk, & Rose, 1987), and a brief background questionnaire. Next, traditional and CogState assessments were completed. In order to control for order effects, the order of traditional and CogState test presentation was counterbalanced so that CogState presentation occurred first for some participants and second for others (Shadish et al., 2002). Appendix H shows the order with which tasks were presented, as well as administration times for each.

Prior to CogState test administration, participants were first provided a scripted tutorial of the program and laptop workings (see Appendix I). Testing began following verbalized understanding of computer workings and demonstrated competency with mouse use. Traditional test administration measures likewise began with formal instructions. Total test participation took approximately 1.5 hours for each participant to complete.

When both sets of tests were completed, participants were orally debriefed and provided information about the study and support group information on a paper handout (Appendix J).
Measures

Demographics and Medical Background

The survivor and partner each completed brief questionnaires (see Appendices E-F) regarding information about their socio-economic status, education, employment, substance use, and sleep behavior. HCT survivors completed six additional questions about their transplantation and prior treatment regimens.

Psychopathology

Anxiety, depression, and fatigue are known confounds of neuropsychological impairment (Lezak, 1983). As such, the following assessments were administered as a brief screening of these constructs.

**PHQ-2.** (Kroenke et al., 2003; See Appendix G). The PHQ-2 is a brief, two-question measure of depression. It inquires about the frequency of depressed mood and anhedonia over the past two weeks, scoring each as 0 (“not at all”) to 3 (“nearly every day”). The total score can range from 0 to 6. With use of a cut-point score of ≥ 3, the PHQ-2 has demonstrated high sensitivity (.83) and specificity (.90) for detecting major depression. It has shown lower sensitivity (.62) for predicting other depressive disorders, but higher specificity (.95).

**GAD-2.** (Kroenke et al., 2007; See Appendix G). The GAD-2 is a brief anxiety screening measure composed of the two core criteria for generalized anxiety disorder, which have also been shown to be good screening items for other anxiety disorders. Each of the two items is rated on a scale of 0 (“not at all”) to 3 (“nearly every day”). The total score may range from 0 to 6. A cut-point of 3 or greater has shown high sensitivity in diagnosing generalized anxiety disorder (.86), panic disorder (.76), social anxiety disorder (.70), and posttraumatic stress disorder (.59) as well as high specificity (.81-.83) within these disorders as well.
Sleep

**SPQ.** (Jenkins et al., 1987; See Appendix G). The SPQ is a brief sleep scale designed to assess sleep disturbance, consisting of four items for evaluating the prevalence of sleep difficulties over a one-month period. Items are answered on a scale of 0 (not at all) to 5. Higher scores are indicative of greater sleep difficulty. The SPQ has shown strong internal consistency reliability (Cronbach’s $a = 0.79$) and likewise demonstrates temporal stability ($r_{tt} = .42-.59, p < 0.001$).

Substance Abuse

**The CAGE Questionnaire.** (Ewing, 1984). The CAGE is a four-item alcohol screener used to assess alcohol use. The four questions are based on Cutting down, Annoyance by criticism, Guilty feeling, and Eye-openers. CAGE scores $\geq$ two have shown specificity (76%) and sensitivity (93%) as well as for identifying alcoholism (77% and 91% respectively).

Premorbid Intelligence

**North American Adult Reading Test (NAART).** (Nelson, 1982) The NAART assesses premorbid intelligence. It is composed of 50 irregularly spelled words, such as “debt” or “naïve.” Words are printed in two columns on both sides of an 8½” x 11” card. This card is given to the subject to read, and the examiner records the errors. The time required for administration is listed as approximately ten minutes (Spreen & Strauss, 1991), although study participants completed the assessment in less than five minutes. Each incorrectly pronounced word counts as one error, though variations due to regional accents are considered to be acceptable. The total number of errors are tabulated and entered into formulas to compute estimated scores of Verbal (VIQ), Performance (PIQ) and Full Scale (FSIQ) IQs. For VIQ and FSIQ, a positive discrepancy of 15 or more points between estimated and actual IQ scores indicates the possibility of intellectual
deterioration or impairment (based on the calculation of 95% confidence levels), while for PIQ, a positive discrepancy of at least 21 points between estimated and actual IQs indicates the possibility of deterioration (Spreen & Strauss, 1998). The NAART possesses high inter-rater reliability (.96-.98) as well as high test-retest reliability (.98; Crawford, Park, Stewart, Besson & De Lacey, 1989). Crawford et al. (2001) reported a correlation of .77 between childhood (age 11) IQ test scores and NAART scores in adults at the age of 77. Early studies furthermore indicated high correlation between performances on the NAART with the general factor of intelligence “g” from the Wechsler scales (Crawford et al., 1989).

**Neuropsychological Functioning**

**CogState tasks.** The proposed Adapted CogState Brief Battery consisted of seven CogState tasks: Detect, Identify, One-Back, Learn, CPAL, GMLT, and ISLT. The first four of these tasks, Detect, Identify, One-Back, and Learn, make up the CogState Brief Battery and respectively assess processing/psychomotor speed, visual attention, working memory, and visual learning, and memory. These tasks were chosen because they measure neuropsychological domains which have been shown to be impaired in HCT patients. This set of tasks has been examined in a prior validation study by CogState authors (Maruff et al., 2009), after which this study is modeled.

Given that the CogState Brief Battery does not measure executive function and verbal learning/memory, it was expanded to include measures of these domains, since deficits in both areas have been documented in HCT survivors. The GMLT and ISLT, measures of executive function and verbal learning/memory, respectively, from the full CogState battery set, were included in order to provide a more comprehensive assessment of cognitive difficulties specific
to HCT patients. CPAL, an additional CogState measure of visual memory and learning, was likewise included for completeness.

All CogState tasks were accessed via software available for download via Internet. Testing was completed via computer, in the presence of a test administrator, who provided start/end times and brief verbal prompts for tasks. Assessment data were then uploaded online to the CogState data capture system, DataPoint.

DataPoint is accessible online; from the secure website, clients can download comprehensive performance reports, which include raw data, a summary data file, and a “clinical style report.” The CogState raw data report provides basic keystroke and computer-generated time point information. The summary data set consists of participants’ test results according to task as well as descriptive statistics and accuracy counts. The set excludes “post-anticipatory correct responses” or premature responses “not seen to be reflective of an actual calculated response.” Following post-anticipatory correct responses, subjects tend to respond more slowly than to the stimulus directly, due to having to disengage before re-engage to the stimulus. CogState tasks seek out the post-anticipatory correct responses and add extra trials, in order to measure a full set of correct trial responses, which are provided in the summary data set (D. Darby, personal communication, June, 14, 2011; See Appendix L). Transformations of the speed and accuracy data (i.e., the log 10 of each response time) are also included within the summary data set. Finally, “clinical style reports” are provided for Detection, Identification, One back, Learn and GMLT errors; these include individual standardized scores ($M = 100, SD = 10$) and standardized change scores for subjects, based on normative data. For this study, data from participant data summary files were used in analyses.
**Processing/psychomotor speed.** The following CogState task provided a measure of processing/psychomotor speed.

**Detect.** This task takes two minutes to administer and also targets attention and psychomotor function. It uses a playing card scheme, requiring that the participant to identify when a card has been turned over. The task begins with a playing card presented in the center of the screen. The card will flip over so it is face up and when it does, the participant must press the “yes” key, denoted as “k” for right-handed persons or “d” for the left-handed. The card then goes to the “back of the pack” and the participant must press the “yes” key as soon as the next card flips over, and so forth. Responses are measured in Log10 milliseconds, with an average of correct responses providing the full task score. Lower scores (i.e., quicker response times) on this task are indicative of better performance. The Detection test has demonstrated strong test-retest reliability (.76) over an interval of 1.5 hours (Collie, 2003). It has also been shown to moderately correlate with TMT-A (.60), TMT-B (.59), and with the GPB (Kløve, 1964), dominant hand task (.57; Cysique, 2006). Divergent validity has also been established as the Identification Task was found to negatively correlate with the CVLT (Delis et al., 1996; r = -.28; Cysique, 2006). Temporal stability has been demonstrated over both short intervals (1–4 hours) as well as over a thirty-day delay, demonstrating only non-significant score changes (Pietrzak et al., 2009).

**Visual attention.** The following CogState task provided a measure of visual attention.

**Identify.** Identify, a measure of attention, presents a series of playing cards and asks the participant to identify if the card is red, using keyboard buttons “k” and “d,” respectively, for “yes” or “no.” For left-handed participants, these settings can be reversed. The task is presented for 2 minutes. Responses are measured in Log10 milliseconds, and the speed of performance is calculated for the mean of the log 10 transformed reaction times for correct responses. Lower
scores (i.e., quicker response times) on this task are indicative of better performance. Test-retest reliability within this task is modest (.61) for a test-retest interval of 1.5 hours (Collie, 2003). The Identification task has also been shown to moderately correlate with TMT-A (.60), TMT-B (.49), and with the GPB (Kløve, 1964), dominant hand task (.62; Cysique, 2006). Divergent validity has also been demonstrated by a negative correlation (-.31) between the Identification Task and CVLT (Delis et al., 1996; Cysique, 2006). Temporal stability has been demonstrated over both short intervals (1-4 hours) as well as over a thirty-day delay, demonstrating only non-significant score changes (Pietrzak et al., 2009).

**Working memory.** The following CogState task provided a measure of working memory.

**One-Back.** One-Back is a measure of working memory and attention. It also uses a playing card scheme and requires that the participant attend to a card presented on the center of the screen, and answer the question: “Is this card the same as that on the immediately previous trial?” The participant is required to respond using keyboard buttons “k” and “d,” respectively, for “yes” or “no.” As with other CogState tasks, these settings may be reversed for left-handed participants. The One-Back task is presented for three minutes or until 42 trials have been completed. Trial performance is represented via arcsine proportion of correct responses; the accuracy of participant performance is measured via arcsine transformation of the square root of correct response proportion. A higher score on this task is indicative of better performance. The One-Back task has shown significant ($p < .01$) correlations to TMT-A (Reitan & Wolfson, 1985; .69), TMT-B (Reitan & Wolfson, 1985; .71), SDMT (.81), Span (Wechsler, 1997; .80), and BVMT (.54; Maruff et al., 2009).
Visual learning and memory. The following CogState tasks provided measurement of visual learning and memory.

Learn. Learn is a measure of visual learning and memory which takes five minutes to administer. A playing card is presented in the center of the computer screen, and the participant must respond as to whether he or she has seen the card before in the task. A new card is immediately presented following each response, and the participant must decide for each card if he has previously been shown the card, using the keyboard buttons “k” and “d,” respectively, for “yes” or “no.” As with other CogState tasks, these settings may be reversed for left-handed participants. If an incorrect response is given, an error noise is heard. Trial performance is represented via arcsine proportion of correct responses; the accuracy of participant performance is measured via arcsine transformation of the square root of correct response proportion. A higher score on this task is indicative of better performance. The Learn task has shown significant correlations ($p < .01$) to TMT-B (Reitan & Wolfson, 1985; .59), SDMT (.57), Span (Wechsler, 1997; .69), BVMT (.83) and RCFT-R (Osterrieth, 1944; .79; Maruff et al., 2009). Temporal stability has been demonstrated over both short intervals (1-4 hours) as well as over a thirty-day delay, demonstrating only non-significant score changes (Pietrzak et al., 2009).

Continuous Paired Associate Learning Task (CPAL). The CPAL is a measure of visual learning and memory. It uses a simulated ball scheme, composed of nine balls positioned on the screen; the center ball is the target object. The full task is comprised of two stages. At the outset of stage 1, the participant is prompted with on-screen instructions asking, “In what locations do these pictures belong?” They must learn and remember the pictures hidden beneath different ball locations on the screen. Participants first tap the central ball to begin. Each picture to be learned will be revealed from this central ball location. As each picture is revealed, the subject must tap
the other ball location and remember where that matching picture was located. In stage two of the task, participants are prompted with the same on-screen instructions asking, “In what locations do these pictures belong?” Then the pictures will be revealed sequentially underneath the center ball. As they appear, the subject must tap on the peripheral location where that picture previously appeared. The total score of this task is represented by the number of errors; thus, a lower score is indicative of better performance. Reliability and validity of the Continuous Paired Associate Learning Task has yet to be established.

**Executive function.** The following CogState task provided a measure of executive function.

**Groton Maze Learning (GML).** GML is a measure of executive function and takes a total of 5 minutes to administer. The participant is shown a 10 x 10 grid of tiles overlying a hidden 28-step pathway among the 100 possible tile locations. The start of the pathway is indicated by a blue tile at the top left of the screen, while the finish location is indicated by a tile with red circles at the bottom right of the screen. The participant is instructed to move one step from the start location and to continue, one tile at a time, toward the end at the bottom right. The subject moves by touching the tile next to their current location with a mouse. After each move, the computer indicates a correct response with a green check mark, symbolizing the next step in the pathway. An incorrect response is denoted with a red cross; the participant must then go back to the last correct location and make a different tile choice to advance toward the pathway end. Participants are presented with two rules: they cannot move diagonally or backwards. Attempts to do so will be recorded as rule breaks, while choosing a tile not part of the hidden pathway is recorded as an error. Once the participant has correctly navigated the 28-step pathway, this task (with the same pathway) is repeated 4 times. The total task score is based upon the number of
errors made in attempting to learn the same hidden pathway on five consecutive trials at a single session. A lower score (i.e., fewer errors) on this task is indicative of better performance. There are 20 alternate forms of the GML, within which different pathways are presented.

The GML has shown moderate relationships with other CogState measures including Detection (.32), a measure of attention, and Monitoring (.40), a measure of visual attention (Pietrzak, 2008). Furthermore, a moderate correlation (.59) has been shown between the GMLT and the Paced Auditory Serial Addition test (PASAT), a measure of sustained attention, concentration, and speed of processing. Internal consistency reliability of the GMLT has likewise been demonstrated. The total error score within this task is comprised of legal errors, perseverative errors, and rule-break errors. Strong correlations between legal errors and perseverative errors (.99) and rule-break errors (.73) have been found, as well as between rule-break errors and perseverative errors (.77). Legal errors furthermore were found to correlate strongly (.60) with another measure of planning, the Tower of Toronto (TOT; Pietrzak, 2007). Temporal stability has been demonstrated over both short intervals (1-4 hours) as well as over a thirty-day delay, demonstrating only non-significant score changes (Pietrzak et al., 2009).

**Verbal learning and memory.** The following CogState task provided a measure of verbal learning and memory.

**International Shopping List (ISL).** The ISL takes five minutes to administer, during which the test administrator reads a shopping list. At the end of the word presentation, the participant is asked to remember as many items as he/she can. As the participant names a listed item, the test supervisor clicks a button on the screen with the mouse. If the participant says a word that was not on the list, the test supervisor will click “other word.” Repeated words are likewise recorded, and an “undo” button is available for supervisor mistakes. Once the
participant can no longer recall any more items, the test is repeated two more times. A delayed recall trial is also administered, where, fifteen minutes later, the participant is asked to remember as many words as she/he can from the previous trial administrations. The total score of each task is represented by the number of correct responses; thus, a higher score is indicative of better performance. The ISL correlates strongly with the HVLT-R \((r > .8; \text{Brandt \\& Benedict, 2001})\). ISL has shown statistically significant \((p < .01)\) test-retest reliability the first three test trials \((r = .56 - .73, )\), as well as for the delayed recall trial \((r = .45; \text{Thompson et al., 2011})\).

**Traditional Neuropsychological Tests**

A series of traditional neuropsychological tests were administered. A review of these tests follows.

**Processing/psychomotor speed.** The following traditional neuropsychological tests provided measurement of attention.

*Grooved Pegboard Dominant and Non-Dominant (GPB-D, GPB-ND).* (Kløve, 1964). The GPB is a measure of motor speed and dexterity. It requires that subjects place a peg into a board containing a 5 x 5 set of slotted holes angled in different directions. Each page has a ridge along one side and requires rotation for correct insertion into the hole. Subjects are timed on completion of this task with both dominant and dominant hands. Scores are reported in terms of time to completion in seconds. Higher scores are indicative of worse performance. Standardized scores are age and education-normed, and the test takes approximately two minutes to administer. The GPB demonstrates a moderate to strong temporal stability, yielding test-retest correlation of .67-.86 (Dikmen, Heaton, Grant, & Temkin, 1999; Levine, Miller, Becker, Selnes & Cohen, 2004).
**Trail Making Test A (TMT-A).** (Reitan & Wolfson, 1985). The TMT-A is one of two parts of the full TMT, which together assess visual scanning, complex attention, psychomotor speed, and mental flexibility. The TMT-A requires that 25 encircled numbers, randomly arranged on a page, be connected in proper order. The test takes approximately two minutes to administer, and scoring is expressed in terms of the time taken for both parts of the test. Psychometrics of the full TMT are reviewed below.

**Visual attention.** The following traditional neuropsychological tests provided measurement of visual attention.

**Trail Making Test A&B (TMT).** (Reitan & Wolfson, 1985). As previously noted, TMT assesses a variety of domains, including visual scanning, complex attention, psychomotor speed, and mental flexibility. The two TMT parts, TMT-A and TMT-B, each require subjects to connect points, via drawing, on a sheet of paper. TMT-A requires encircled numbers to be connected in proper order, whereas TMT-B requires that 25 randomly arranged (and encircled) numbers and letters be connected in alternating, ascending order. The full test takes approximately five minutes to administer, and scoring is expressed in terms of the time taken for both parts of the test. Test-retest reliability of the full TMT has been extensively reviewed by Spreen and Strauss (1998). Coefficients between initial administration and one-year post were reported to range from .69-.94 for TMT-A and .66-.86 for TMT-B (Goldstein & Watson, 1989). Lower test-retest reliability for TMT-A (.55) was reported after a 3-week interval but was adequate for TMT-B (.75; Bornstein, 1987). Strong interrater reliability has also been reported for the both TMT-A (.94) and TMT-B (.90; Fals-Stewart, 1991).

**Symbol Digit Modalities Test (SDMT).** (Smith, 1982). The SDMT is a test requiring visual processing and divided attention. It is printed in a test booklet and consists of four rows
containing 100 small blank squares. Each square is paired with a with a nonsense symbol. Above these rows is a printed key that pairs each number with a different symbol. After undergoing a practice trial, subjects are to fill in the blank squares with the number that is paired to the number above the blank space as quickly as possible. The task continues for 90 seconds. High scores are indicative of better performance. The SDMT correlates strongly with the Digit Symbol test ($r = .80$) and likewise possesses strong test-retest reliability ($r = .80$; Lezak, 2006; Wechsler, 1997). It has shown strong criterion validity in a comparison of 100 brain lesion patient scores to normative data, correctly identifying 86% of the patient group using a cut-point of 1.5 standard deviations below the age norm (Lezak, 2006).

**Working memory.** The following traditional neuropsychological tests provided measurement of working memory.

*Symbol Digit Modalities Test (SDMT).* (Smith, 1982). See above description.


*Wechsler Memory Scale-III Spatial Span Subtest* (**Span**). (Wechsler, 1997). The Span subtest is a measure of visual scanning and working memory, taking approximately five minutes to administer. It is a visual analogue of the Digit Span test, consisting of two parts: Spatial Span Forward and Spatial Span Backward. The subject is presented with a series of cubes. For each part, the test administrator taps a series of cubes at the rate of about one cube per second. Following this presentation, the subject either taps the cubes in the same order as the examiner (Spatial Span Forward) or in reversed order (Spatial Span Backward). The first trial of this test begins with two cubes and continues to include eight cubes. Subjects are given two trials at each series length, and testing continues until both trials of the series are failed. One point is awarded for each correctly answered trial. The maximum score for the test is 32; 16 points may be
awarded for each part. Of other WMS-III tests, Span has illustrated the strongest correlation with the letter-number sequencing subtest (.45). Notably interscorer agreement is high for this task, on average in the high .90s; Span is susceptible to practice effects ($r_{n} = .69$).

**Visual learning and memory.** The following traditional neuropsychological tests provided measurement of visual learning and memory.

*Rey Complex Figure Test-Delayed Recall (RCFT-R).* (Osterrieth, 1944). The RCFT-R is a test of a variety of cognitive processes, including planning and organizational skills, problem-solving strategies, perceptual, motor and memory functions (Waber & Holmes, 1986). Materials include blank pieces of paper, colored pencils, and the Rey-Osterrieth figure. The procedure requires that the subject copy the figure, and the task is timed. It takes approximately ten minutes to administer, excluding the thirty-minute delay, at which time the subject is instructed to reproduce the figure from memory. The subject has unlimited time to reconstruct the figure. The test is scored based on a rubric outlining 18 scorable elements, with .5-2.0 points being rewarded for each element, depending on accuracy/distortion and location of reproduction. The RCFT-R possesses strong inter-rater reliability (.93-.99) as well as temporal stability ($r_{t} = .888$). Further, it illustrates robust construct validity, statistically significantly correlating the strongest with WAIS-R (Wechsler, 1987) performance subtests including digit symbol ($r_{p} = .533, p < .001$) and block design ($r_{p} = .451, p < .01$).

**Executive function.** The following traditional neuropsychological tests provided measurement of executive function.


*Symbol Digit Modalities Test (SDMT).* (Smith, 1982). See above description.
Verbal learning and memory. The following traditional neuropsychological tests provided measurement of verbal learning and memory.

Hopkins Verbal Learning Test- Revised (HVLT-R), (Benedict et al., 1998). The HVLT-R assesses verbal learning and memory. There are six alternate forms of the HVLT-R, each of which is composed of a list of 12 nouns. The lists are made up of 3 semantic categories (e.g. four-legged animals or human dwellings), with 4 of the 12 listed nouns fitting into each of these categories. The semantic categories are different across the six forms of the HVLT-R. Each of the HVLT-R forms includes three learning trials, a delayed recall trial (given after a delay of 20-25 minutes without forewarning), and a yes/no recognition trial. The recognition trial is composed of a list of 12 “target” and 12 “non-target” words, all random. Of these words, six are drawn from the same semantic categories as the targets. Administration takes approximately five minutes during which the tester reads the word list, asking the subject to repeat the list, in any order, both immediately and after a delay. Recall performance is then recorded, the maximum score for each recall trial being 12. The recall scores are combined to form three measures of learning and memory: The Total Recall score, % Retention, and Recognition Discrimination Index. The Total Recall score is the sum of learning trials 1-3. The % Retention, based on the percent of words retained after the delay, is calculated by dividing trial 4 recall by the best of Trials 2 and 3. The Recognition Discrimination Index is the number of true positives minus false positives. The HVLT-R has demonstrated moderate-strong temporal stability reliability. Test-retest reliability coefficients calculated for the four primary HVLT-R variables following a mean interval of six weeks were .74 for Total Recall, .66 for Delayed Recall, .39 for % Retention, and .40 for The Recognition Discrimination Index (Benedict et al., 1998). Convergent validity of the HVLT-R has also received modest support. The following correlations were found between the
HVLT-R and CVLT (Delis et al., 1996): total learning ($r = .36$), delayed recall ($r = .62$), intrusion errors ($r = .34$), and recognition hits ($r = .48$; Lacritz, 2001). Furthermore, correlations between the HVLT-R and Logical Memory (Wechsler, 1987; $r = .65-.77$) and Visual Reproduction (Wechsler, 1987; $r = .54-.69$) indicate that the test correlates more strongly with verbal memory than visual memory (Shapiro, 1999).

**Data Analyses**

**Design**

A cross-sectional design was used as it allowed construct and criterion validity to be evaluated. Construct validity was evaluated through comparing scores of control subjects on CogState with traditional neuropsychological measures. Criterion validity was evaluated through examining the ability of CogState tasks to discriminate HCT survivors from control subjects.

**Statistical Analyses**

All statistical analyses were computed using SPSS 20 and SPSS 21. The first set of analyses addressed hypotheses associated with CogState construct validity, and the second set addressed criterion validity.

**Construct validity.** Pearson product-moment correlations were computed between CogState measures and comparator neuropsychological tests in HCT survivors as indicated in Table 5, which presents these tests according to cognitive domain.
In order to more thoroughly assess the magnitude of associations, statistically significant correlation coefficients were then transformed into effect sizes. Corresponding confidence intervals for each effect size were calculated, plotted, and examined for overlap, as non-overlapping confidence intervals are indicative of statistically significant differences (Cumming & Finch, 2005).

An additional series of exploratory analyses were conducted within which effect sizes were computed and graphed, with 95% CIs for all correlation coefficients that reflected a medium effect size (e.g., $r = .30$).

**Criterion validity.** The hypotheses addressing the criterion validity of the Adapted CogState Brief Battery were addressed by evaluating the magnitude of HCT survivors’ impairment on CogState as compared with healthy control subjects.
An ANCOVA was performed to assess HCT survivor performance on each CogState task, relative to matched control subjects, after accounting for variation due to premorbid intelligence. The magnitudes of group differences (i.e., $F$ statistics) were transformed into effect sizes (e.g., $d$), and the non-overlap statistic ($\%NOL$) was computed for each difference. The $\%NOL$ provides a measurement of score overlap between control and clinical groups and was beneficial for determining the extent to which HCT survivor performance differed from control group performance. Greater $\%NOL$ reflects the proportion of the clinical group whose performance is not shared by the control group and is thus indicative of better classification. A $\%NOL$ of 93% is indicative of a good clinical marker (Zakzanis, 2000).

The $\%NOL$ is derived from the overlap statistic ($\%OL$) and has been recommended for use in evaluating validity of neuropsychological test batteries by Zakzanis (2001). The $\%OL$, seen to represent the amount of test measure overlap between clinical and control samples, reflects the extent to which the distribution of clinical and control sample scores overlap. It is calculated through the use of Cohen’s table of non-overlap values (1988). The table presents overlap values for each listed effect size; 100% overlap corresponds with an effect size ($d$) of 0. The $\%N-OL$, proposed for use within this study, is calculated by subtracting the $OL\%$ from 100 (Zakzanis, 2001).

In order to control for Type I error rates, (i.e., inflating the chance of significance when it is not actually present), a Bonferroni correction was utilized. This correction allowed for family-wise error to be controlled, correcting the level of significance for each test so that the overall Type I error rate ($a$) remained at .05 across all comparisons.

A series of exploratory analyses were also computed, examining group differences without accounting for variation due to pre-morbid intelligence, through a series of t-tests.
Group differences (i.e., $t$ statistics) were transformed into effect sizes (i.e., $d$), and %NOL was computed. These analyses replicate the analyses in Maruff et al.’s (2009) validation study.

**Results**

**Psychopathology Prevalence**

Criteria for major depression was met by 12.5% of the entire study sample; 15% of survivors met criteria and 10% of spousal partners met criteria for major depression.

Of the whole sample, 7.5% met criteria for clinically significant anxiety as screened by the GAD-2; 5% of survivors, 10% of spousal partners. No participants met criteria for alcoholism.

No significant differences with regard to psychopathology measures emerged between groups, although the differences associated with sleep (as per the Sleep Problems Questionnaire (SPQ) approached significance, $t = 1.80, p = .08$, with survivors experiencing more sleep problems. See Table 6 for a summary of participant psychopathology.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Spousal partners</th>
<th>HCT survivors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$ (SD)</td>
<td>Range</td>
</tr>
<tr>
<td>Patient Health Questionnaire-2 (PHQ-2)</td>
<td>0.70 (1.34)</td>
<td>0 – 5</td>
</tr>
<tr>
<td>Generalized Anxiety Disorder-2 (GAD-2)</td>
<td>1.05 (1.27)</td>
<td>0 – 4</td>
</tr>
<tr>
<td>Sleep Problems Questionnaire (SPQ)</td>
<td>6.00 (5.89)</td>
<td>0 – 20</td>
</tr>
<tr>
<td>CAGE Questionnaire (CAGE)</td>
<td>0.10 (0.44)</td>
<td>0 – 2</td>
</tr>
</tbody>
</table>

Means and standard deviations were calculated for HCT survivors’ and spousal partners’ performances on all traditional neuropsychological measures and CogState tasks.

Using impairment criterion of 1.5 standard deviations below the mean on traditional tests, 0-40% of HCT survivors exhibited impaired performance across traditional neuropsychological measures. Zero to 30% of spousal partners exhibited impaired performance.
Using impairment criterion of 2.0 standard deviations below the mean on traditional tests, 0-40% of HCT survivors exhibited impaired performance across tests, as did 0-10% of spousal partners. See Table 7 for a summary of these findings. At the time of data analyses, normative data for CPAL was only available for ages 18 – 60; as such, standard scores for six participants could not be conducted. Numbers in Table 7 reflect data for 20 HCT survivors and 20 spousal partners on all CogState tasks except CPAL, for which 17 HCT survivors and 17 spousal partners’ data are reflected.

Table 7

*Traditional test z-score means, standard deviations, and percentages of HCT survivor and spousal partner impairment*

<table>
<thead>
<tr>
<th>Domain</th>
<th>Test</th>
<th>M</th>
<th>SD</th>
<th>n (%) impaired ≤ -1.0 SD</th>
<th>≤ -1.5 SD</th>
<th>≤ -2.0 SD</th>
<th>M</th>
<th>SD</th>
<th>n (%) impaired ≤ -1.0 SD</th>
<th>≤ -1.5 SD</th>
<th>≤ -2.0 SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processing/ Psychomotor Speed</td>
<td>GPB-D</td>
<td>-1.43</td>
<td>1.16</td>
<td>14 (70)</td>
<td>8 (40)</td>
<td>8 (40)</td>
<td>-0.75</td>
<td>1.11</td>
<td>10 (50)</td>
<td>4 (20)</td>
<td>3 (15)</td>
</tr>
<tr>
<td></td>
<td>GPB-ND</td>
<td>-1.16</td>
<td>1.05</td>
<td>12 (60)</td>
<td>7 (35)</td>
<td>4 (20)</td>
<td>-1.00</td>
<td>0.89</td>
<td>9</td>
<td>6 (30)</td>
<td>3 (15)</td>
</tr>
<tr>
<td>Visual Attention</td>
<td>TMT-A</td>
<td>-0.33</td>
<td>0.73</td>
<td>5 (25)</td>
<td>0</td>
<td>0</td>
<td>-0.10</td>
<td>0.74</td>
<td>2 (10)</td>
<td>1 (5)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Working Memory</td>
<td>SDMT</td>
<td>-0.09</td>
<td>0.85</td>
<td>4 (20)</td>
<td>1 (5)</td>
<td>0</td>
<td>0.41</td>
<td>0.85</td>
<td>1 (5)</td>
<td>1 (5)</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Span</td>
<td>0.40</td>
<td>0.78</td>
<td>2 (10)</td>
<td>0</td>
<td>0</td>
<td>0.57</td>
<td>0.79</td>
<td>1 (5)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Visual Learning/ Memory</td>
<td>RCFT</td>
<td>-1.04</td>
<td>1.24</td>
<td>9 (45)</td>
<td>7 (35)</td>
<td>5 (25)</td>
<td>-0.88</td>
<td>1.15</td>
<td>8 (40)</td>
<td>4 (20)</td>
<td>4 (20)</td>
</tr>
<tr>
<td>Executive Function</td>
<td>TMT-B</td>
<td>-0.33</td>
<td>1.06</td>
<td>7 (35)</td>
<td>3 (15)</td>
<td>0</td>
<td>0.37</td>
<td>0.95</td>
<td>1 (5)</td>
<td>0</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Verbal Learning/ Memory</td>
<td>HVLT</td>
<td>-0.12</td>
<td>0.95</td>
<td>5 (25)</td>
<td>1 (5)</td>
<td>0</td>
<td>-0.54</td>
<td>1.10</td>
<td>8 (40)</td>
<td>4 (20)</td>
<td>2 (10)</td>
</tr>
<tr>
<td></td>
<td>HVLT-R</td>
<td>0.01</td>
<td>1.62</td>
<td>6 (30)</td>
<td>2 (10)</td>
<td>2 (10)</td>
<td>-0.56</td>
<td>0.97</td>
<td>10 (50)</td>
<td>3 (15)</td>
<td>2 (10)</td>
</tr>
<tr>
<td></td>
<td>HVLT-RET</td>
<td>-0.24</td>
<td>1.03</td>
<td>4 (20)</td>
<td>3 (15)</td>
<td>2 (10)</td>
<td>-0.23</td>
<td>0.72</td>
<td>5 (25)</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Using impairment criterion of 1.5 standard deviations below the mean on CogState tasks, 0-15% of HCT survivors exhibited impaired performance across tasks, as compared with 0-10% of spousal partners.

Using impairment criterion of 2.0 standard deviations below the mean on CogState tasks, 0-5% of HCT survivors and spousal partners alike exhibited impaired performance across tasks.

See Table 8 for a summary of these results.

Table 8

<table>
<thead>
<tr>
<th>CogState task z-score means, standard deviations, and percentages of HCT survivor and spousal partner impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>a.</td>
</tr>
<tr>
<td>b.</td>
</tr>
<tr>
<td>c.</td>
</tr>
<tr>
<td>d.</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>e.</td>
</tr>
<tr>
<td>f.</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Prior to computing the primary analyses, data were examined for outliers. One participant evidenced a broad range of CogState scores (One-Back $z = -3.34$; CPAL $z = 3.24$). In order to investigate the effects of this potential outlier, analyses were computed two ways: both including and excluding this participant. Exclusion of scores did not change results, so this participant’s scores were kept in the dataset.

**Construct Validation**

**Correlational analyses.** In order to test the first set of hypotheses addressing CogState construct validity, Pearson product-moment correlations were computed. As illustrated in Table 9, relationships between CogState and traditional tests measuring the same domain were overall weaker than expected; however, a number of interesting relationships emerged.

<table>
<thead>
<tr>
<th></th>
<th>Detect</th>
<th>Identify</th>
<th>One-Back</th>
<th>Learn</th>
<th>CPAL</th>
<th>GML</th>
<th>ISL</th>
<th>ISL-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPB-D</td>
<td>-0.29</td>
<td>-0.39</td>
<td>0.05</td>
<td></td>
<td>0.67**</td>
<td>0.34</td>
<td>0.31</td>
<td>-0.01</td>
</tr>
<tr>
<td>GPB-ND</td>
<td>-0.26</td>
<td>-0.21</td>
<td>-0.01</td>
<td></td>
<td>0.44*</td>
<td>0.31</td>
<td>0.27</td>
<td>-0.27</td>
</tr>
<tr>
<td>TMT-A</td>
<td>-0.49*</td>
<td>-0.42</td>
<td>-0.46*</td>
<td></td>
<td>0.10</td>
<td>0.10</td>
<td>-0.02</td>
<td>-0.20</td>
</tr>
<tr>
<td>SDMT</td>
<td>0.04</td>
<td>-0.18</td>
<td>0.21</td>
<td>0.38</td>
<td></td>
<td>0.31</td>
<td>0.11</td>
<td>0.17</td>
</tr>
<tr>
<td>Span</td>
<td>-0.07</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.39</td>
<td></td>
<td>0.15</td>
<td>-0.33</td>
<td>-0.12</td>
</tr>
<tr>
<td>RCFT</td>
<td>0.23</td>
<td>0.33</td>
<td>0.09</td>
<td>0.20</td>
<td></td>
<td>-0.42</td>
<td>-0.22</td>
<td>0.06</td>
</tr>
<tr>
<td>TMT-B</td>
<td>-0.05</td>
<td>-0.28</td>
<td>-0.08</td>
<td>0.28</td>
<td></td>
<td>-0.48*</td>
<td>-0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>HVLT</td>
<td>-0.05</td>
<td>-0.05</td>
<td>0.44*</td>
<td>0.14</td>
<td></td>
<td>0.01</td>
<td>0.05</td>
<td>0.35</td>
</tr>
<tr>
<td>HVLT-R</td>
<td>-0.07</td>
<td>0.00</td>
<td>0.29</td>
<td>-0.01</td>
<td></td>
<td>0.25</td>
<td>-0.17</td>
<td>0.13</td>
</tr>
<tr>
<td>HVLT-RET</td>
<td>0.18</td>
<td>0.24</td>
<td>0.43*</td>
<td>0.09</td>
<td></td>
<td>0.09</td>
<td>-0.18</td>
<td>0.64**</td>
</tr>
</tbody>
</table>

* $p < .05$; ** $p < .01$
**Processing/Psychomotor speed.** Detect, a CogState measure of processing/psychomotor speed, was hypothesized to correlate most strongly with traditional tests of this domain, including GPB and TMT-A. This hypothesis was supported. The measure did exhibit a strong negative correlation with TMT-A \( (r = -.49, p < .05) \) as expected. Detect associations with GPB-D \( (r = -.29) \), and GPB-ND \( (r = -.26) \) were also negatively correlated as expected.

**Visual attention.** Identify, a CogState measure of visual attention, was hypothesized to correlate most strongly with traditional tests of this domain, including TMT-A, TMT-B, and SDMT. It exhibited strong relationships with TMT-A \( (r = -.42, p = .07) \) and B \( (r = -.28; p = .23) \) in the direction expected. Its relationship with SDMT \( (r = -.18; p = .46) \) exhibited a negative correlation as predicted but was far short of significance.

**Working memory.** One-Back, a CogState measure of working memory, did not perform as hypothesized by correlating strongly with traditional measures of working memory (i.e., Span, SDMT) although correlations evidenced were positive, as predicted. Correlations with Span \( (r = .01) \) and SDMT \( (r = -.21) \) were both statistically non-significant. Notably, it significantly correlated with HVLT \( (r = .44, p = .05) \) and TMT-A \( (r = .46, p < .05) \), measures of verbal learning/memory and processing speed/attention, respectively.

**Visual learning and memory.** Learn and CPAL, both CogState measures of visual learning and memory, exhibited no relationship to each other \( (r = -.03) \); they also evidenced relationships with other tests in unexpected directions. Both of the tasks were posited to correlate most strongly with RCFT-R, a traditional of visual learning and memory. Although CPAL illustrated a moderate relationship with RCFT-R \( (r = -.42, p = .10) \), which approached statistical significance, Learn was only mildly correlated with the measure \( (r = .20) \), at a statistically non-significant level. Notably, both were correlated with RCFT in predicted directions.
Interestingly, Learn did correlate significantly with both GPB-D \((r = .67, p = .00)\) and GPB-ND \((r = .44, p = .05)\), measures of processing speed. Furthermore, its relationship with Span, a measure of working memory was stronger and exhibited a trend toward significance \((r = .39, p = .09)\).

CPAL illustrated a strong relationship with TMT-B \((r = -.48, p = .05)\), a traditional measure of executive function, in the correct direction.

**Executive function.** Contrary to study hypotheses, GML, a CogState task of executive function, exhibited non-significant correlations with traditional measures of executive function: SDMT \((r = .11)\) and TMT-B \((r = -.05)\); its relationship with TMT-B was in the predicted direction, whereas its relationship with SDMT was in the opposite direction predicted. GML did show a moderate relationship with Span \((r = -.33, p = .15)\), a measure of working memory that trended towards statistical significance.

**Verbal learning and memory.** Finally, ISL (including both immediate (ISL) and delayed recall (ISL-R) trials), the CogState measure of verbal learning and memory, showed strong relationships with a traditional test of verbal learning and memory test (e.g. HVLT immediate (HVLT), delayed recall (HVLT-R), and retention (HVLT-RET) trials) as predicted. In particular, the ISL-R demonstrated strong relationships in predicted directions with HVLT \((r = .73; p = .001)\), HVLT-R \((r = .56; p = .01)\), and HVLT-RET \((r = .54; p = .01)\). ISL illustrated weaker associations with HVLT tests. Although it was significantly correlated with HVLT-ret \((r = .64, p = .00)\), its relationship with HVLT only trended toward significance \((r = .35, p = .13)\). It showed minimal relationship with HVLT-R \((r = .13, p = .60)\).

Correlations are additionally presented within a multi-trait multi-method matrix. See Table 10.
<table>
<thead>
<tr>
<th>Method 1: CogState</th>
<th>Method 2: Traditional neuropsychological tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. A. Processing/Psychomotor Speed (CET)</td>
<td>1.00</td>
</tr>
<tr>
<td>B. Visual Attention (IDN)</td>
<td>0.61**</td>
</tr>
<tr>
<td>C. Working Memory (CINL)</td>
<td>0.58**</td>
</tr>
<tr>
<td>D.a. Visual Learning and Memory (CICL)</td>
<td>0.07</td>
</tr>
<tr>
<td>D.b. Visual Learning and Memory (CPAL)</td>
<td>- 0.44</td>
</tr>
<tr>
<td>E. Executive Function (GML)</td>
<td>- 0.34</td>
</tr>
<tr>
<td>F. Verbal Learning and Memory (ISL)</td>
<td>0.31</td>
</tr>
<tr>
<td>G. Verbal Learning and Memory - Recall (ISL-R)</td>
<td>0.10</td>
</tr>
</tbody>
</table>

| 2. A.a. Processing/Psychomotor Speed (GPB-D) | - 0.29 | - 0.30 | 0.05 | 0.57** | 0.34 | 0.31 | - 0.01 | - 0.02 | 1.00 |      |     |     |      |     |     |     |     |     |
| A.b. Processing/Psychomotor Speed (GPB-ND) | - 0.26 | - 0.21 | - 0.01 | 0.44 | 0.30 | 0.27 | - 0.27 | - 0.31 | 0.70** | 1.00 |      |     |     |      |     |     |     |     |     |
| B. Visual Attention (TIM-A) | - 0.56* | 0.42 | 0.46* | 0.10 | 0.10 | - 0.02 | - 0.20 | - 0.01 | - 0.08 | - 0.17 | 1.00 |      |     |     |      |     |     |     |     |
| C.a. Working Memory (SDMT) | 0.04 | - 0.13 | 0.21 | 0.38 | 0.31 | 0.11 | 0.17 | 0.39 | 0.14 | 0.12 | 0.20 | 1.00 |      |     |     |      |     |     |
| C.b. Working Memory (Scan) | 0.07 | - 0.01 | 0.05 | 0.19 | 0.31 | 0.11 | 0.17 | 0.39 | 0.14 | 0.12 | 0.20 | 1.00 |      |     |     |      |     |     |
| D. Visual Learning and Memory (RCFT) | - 0.23 | 0.33 | 0.09 | 0.20 | 0.42 | 0.22 | 0.06 | 0.15 | - 0.31 | - 0.43 | 0.15 | 0.09 | 0.31 | 1.00 |      |     |     |
| E. Executive Function (TIM-B) | - 0.05 | - 0.26 | - 0.08 | 0.25 | 0.45* | 0.20 | 0.20 | 0.06 | 0.07 | - 0.03 | 0.38 | 0.56** | 0.21 | 0.35 | 1.00 |      |     |
| F. Verbal Learning and Memory - Recall (HVLT) | - 0.05 | 0.05 | 0.44* | 0.14 | 0.01 | 0.00 | 0.22 | 0.15 | - 0.31 | - 0.43 | 0.15 | 0.09 | 0.31 | 1.00 |      |     |     |
| G. Verbal Learning and Memory - Recall (HVLT-TR) | - 0.07 | 0.06 | 0.25 | - 0.01 | 0.25 | 0.17 | 0.13 | 0.55* | - 0.13 | - 0.50 | - 0.04 | 0.13 | 0.35 | 0.36 | - 0.05 | 0.65** | 1.00 |
| G.a. Verbal Learning and Memory - Retention (HVLT-RET) | 0.18 | 0.24 | 0.45* | 0.09 | 0.08 | 0.18 | 0.64** | 0.54* | - 0.17 | - 0.40 | - 0.22 | 0.04 | 0.26 | 0.23 | - 0.18 | 0.49* | 0.71** | 1.00 |

n = 20

** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)
**Analyses of convergence and divergence patterns.** In general, CogState tasks converged with other CogState tasks designed to assess similar, overlapping constructs and diverged from those designed to assess other distinct constructs. Detect, a measure of processing/psychomotor speed, converged with CogState tasks of visual attention (Identify) and diverged from One-Back, a measure of working memory. It likewise diverged as appropriate from a CogState task of verbal learning and memory (ISL). CogState tasks of visual learning and memory (Learn, CPAL) likewise converged appropriately with ISL, as did One-Back.

Patterns of convergence and divergence between CogState and traditional tasks were mixed. Although CogState task of executive function (GML) exhibited a weak and divergent relationship with SDMT, a traditional task of working memory, it did converge with Span, also a measure of working memory. Relationships between CogState visual learning and memory tasks (i.e., Learn and CPAL) both converged with RCFT-R, a traditional task of visual learning and memory. Detect, a CogState measure of processing/psychomotor speed, appropriately diverged from a traditional task of visual learning and memory (RCFT-R), although its relationship with SDMT, a traditional measure of working memory, did not evidence as strong a divergence as one might expect.

**Effect size analyses.** Statistically significant correlations were next transformed into effect sizes (Cohen’s $d$), using a formula recommended by Zakzanis (2001), i.e. **[formula]**. Use of this formula is necessary in order to calculate an effect size from a correlation coefficient. Ninety-five percent confidence intervals for each effect size were constructed, via a formula recommended by Nakagawa & Cuthill (2007), i.e. $95\% \text{ CI} = ES - 1.96se$ to $ES + 1.96se$, where $se_d = \ldots$. See Table 11.
Table 11

Summary of statistically significant Pearson's product-moment correlations, transformed into effect sizes with confidence intervals

<table>
<thead>
<tr>
<th>CogState Task</th>
<th>Traditional Test</th>
<th>$r$</th>
<th>$d$</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detect</td>
<td>TMT-A</td>
<td>-0.49*</td>
<td>1.12</td>
<td>0.23</td>
<td>2.01</td>
</tr>
<tr>
<td>One-Back</td>
<td>HVLT</td>
<td>0.44*</td>
<td>0.98</td>
<td>0.38</td>
<td>1.58</td>
</tr>
<tr>
<td></td>
<td>TMT-A</td>
<td>-0.46*</td>
<td>1.04</td>
<td>0.17</td>
<td>1.91</td>
</tr>
<tr>
<td>Learn</td>
<td>GPB-D</td>
<td>0.67**</td>
<td>1.81</td>
<td>1.09</td>
<td>2.53</td>
</tr>
<tr>
<td></td>
<td>GPB-ND</td>
<td>0.44*</td>
<td>0.98</td>
<td>0.38</td>
<td>1.58</td>
</tr>
<tr>
<td>CPAL</td>
<td>TMT-B</td>
<td>-0.48*</td>
<td>1.09</td>
<td>0.21</td>
<td>1.97</td>
</tr>
<tr>
<td>ISL</td>
<td>HVLT-ret</td>
<td>0.64**</td>
<td>1.67</td>
<td>0.98</td>
<td>2.36</td>
</tr>
<tr>
<td>ISL-R</td>
<td>HVLT</td>
<td>0.73**</td>
<td>2.14</td>
<td>1.33</td>
<td>2.94</td>
</tr>
<tr>
<td></td>
<td>HVLT-R</td>
<td>0.56*</td>
<td>1.35</td>
<td>0.72</td>
<td>1.98</td>
</tr>
<tr>
<td></td>
<td>HVLT-ret</td>
<td>0.54*</td>
<td>1.28</td>
<td>0.66</td>
<td>1.90</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed)
* Correlation is significant at the 0.05 level (2-tailed)

Effect sizes were then graphed with their 95% confidence intervals (see Fig. 1) in order to more thoroughly assess the magnitude of associations. According to Cumming & Fitch (2005), non-overlapping CIs are seen to be indicative of statistically significant differences. As illustrated within Fig. 1, overlap was present across tests.
Given that associations were not as strong as predicted, a set of exploratory analyses were conducted in order to investigate patterns of confidence interval overlap occurring for all CogState and traditional tests with correlations of $r \geq .3$. According to Cohen (1992), correlation coefficients of $\geq .3$ are indicative of a medium effect; as such, all correlations $\geq .3$, regardless of whether or not they reached statistical significance, were investigated. Relationships meeting this relaxed criteria were transformed into effect sizes, and 95% confidence intervals were constructed, as before. See Table 12 for a summary of these results.
Table 12

Summary of Pearson’s product-moment correlations > 0.30, transformed into effect sizes with confidence intervals

<table>
<thead>
<tr>
<th>CogState Task</th>
<th>Traditional Test</th>
<th>( r )</th>
<th>( D )</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower Limit</td>
</tr>
<tr>
<td>Detect</td>
<td>TMT-A</td>
<td>- 0.49*</td>
<td>1.12</td>
<td>0.23</td>
</tr>
<tr>
<td>Identify</td>
<td>RCFT-R</td>
<td>0.33</td>
<td>0.70</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>TMT-A</td>
<td>- 0.42</td>
<td>0.93</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>GPB-D</td>
<td>- 0.39</td>
<td>0.85</td>
<td>0.02</td>
</tr>
<tr>
<td>One-Back</td>
<td>HVLT</td>
<td>0.44*</td>
<td>0.98</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>HVLT-ret</td>
<td>0.43</td>
<td>0.95</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>TMT-A</td>
<td>- 0.46*</td>
<td>1.04</td>
<td>0.17</td>
</tr>
<tr>
<td>Learn</td>
<td>Span</td>
<td>0.39</td>
<td>0.85</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>SDMT</td>
<td>0.38</td>
<td>0.82</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>GPB-D</td>
<td>0.67**</td>
<td>1.81</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>GPB-ND</td>
<td>0.44*</td>
<td>0.98</td>
<td>0.38</td>
</tr>
<tr>
<td>CPAL</td>
<td>RCFT-R</td>
<td>- 0.42</td>
<td>0.93</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>SDMT</td>
<td>- 0.31</td>
<td>0.65</td>
<td>-0.13</td>
</tr>
<tr>
<td></td>
<td>TMT-B</td>
<td>- 0.48*</td>
<td>1.09</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>GPB-D</td>
<td>0.34</td>
<td>0.72</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>GPB-ND</td>
<td>0.30</td>
<td>0.63</td>
<td>0.04</td>
</tr>
<tr>
<td>GML</td>
<td>Span</td>
<td>- 0.33</td>
<td>0.70</td>
<td>-0.10</td>
</tr>
<tr>
<td></td>
<td>GPB-D</td>
<td>0.31</td>
<td>0.65</td>
<td>0.06</td>
</tr>
<tr>
<td>ISL</td>
<td>HVLT</td>
<td>0.35</td>
<td>0.75</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>HVLT-ret</td>
<td>0.64**</td>
<td>1.67</td>
<td>0.98</td>
</tr>
<tr>
<td>ISL-R</td>
<td>HVLT</td>
<td>0.73**</td>
<td>2.14</td>
<td>1.33</td>
</tr>
<tr>
<td></td>
<td>HVLT-R</td>
<td>0.56*</td>
<td>1.35</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>HVLT-ret</td>
<td>0.54*</td>
<td>1.28</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>SDMT</td>
<td>0.39</td>
<td>0.85</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>GPB-ND</td>
<td>- 0.31</td>
<td>0.65</td>
<td>-0.14</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

Effects sizes and confidence intervals for correlation coefficients meeting this relaxed criterion were plotted and examined for overlap. Figure 2 depicts plotted effect sizes.
Confidence intervals were omitted in this figure, as they showed a great deal of overlap, decreasing figure legibility. See Appendix O, which includes a depiction of both effect size and confidence interval plots for correlation coefficients ≥ .3.

Criterion Validation

**Calculation of the non-overlap statistic (\%NOL).** In order to evaluate criterion validity, or the degree to which CogState tasks could accurately predict group membership of the whole sample (e.g. survivors and spousal partners), a series of ANCOVA’s were performed, evaluating the degree of difference between the two groups after accounting for covariance due to pre-morbid IQ (e.g. NAART). See Table 13 for a summary of these results.
Table 13

Summary of Analysis of Covariance Analyses for CogState Tasks and NAART

<table>
<thead>
<tr>
<th>CogState Task</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detect</td>
<td>1</td>
<td>0.05</td>
<td>0.09</td>
<td>0.77</td>
</tr>
<tr>
<td>Identify</td>
<td>1</td>
<td>3.22</td>
<td>3.32</td>
<td>0.08</td>
</tr>
<tr>
<td>One-Back</td>
<td>1</td>
<td>0.52</td>
<td>0.48</td>
<td>0.5</td>
</tr>
<tr>
<td>Learn</td>
<td>1</td>
<td>0.28</td>
<td>0.4</td>
<td>0.53</td>
</tr>
<tr>
<td>CPAL</td>
<td>1</td>
<td>1.39</td>
<td>0.7</td>
<td>0.41</td>
</tr>
<tr>
<td>GML</td>
<td>1</td>
<td>0.02</td>
<td>0.08</td>
<td>0.78</td>
</tr>
<tr>
<td>ISL</td>
<td>1</td>
<td>0.05</td>
<td>0.04</td>
<td>0.85</td>
</tr>
<tr>
<td>ISL-R</td>
<td>1</td>
<td>0.32</td>
<td>0.38</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Given that no significant differences were found between groups, a series of exploratory analyses were conducted in order to assess other potential confounds. A series of ANCOVAs were also performed, using sleep problems (SPQ) and time since transplant as covariates. Use of these constructs as covariates did not change results, and therefore the originally proposed model (e.g. using premorbid IQ as a covariate) was used for the following analyses.

The magnitude of group differences were then transformed into effect sizes ($d$) through use of the following formula proposed by Zakzanis (2001): $d = \frac{\text{Mean}_{A} - \text{Mean}_{B}}{\hat{\sigma}}$. The percent overlap (%OL) was calculated using values proposed by Cohen (1988); percent overlap was subtracted from 100 in order to obtain the non-overlap statistic (%NOL). See Table 14.
Table 14

<table>
<thead>
<tr>
<th>Task</th>
<th>Survivors</th>
<th>Spouses</th>
<th>F</th>
<th>p</th>
<th>d</th>
<th>OL</th>
<th>%NOL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detect</td>
<td>0.34</td>
<td>0.81</td>
<td>0.40</td>
<td>0.77</td>
<td>0.09</td>
<td>0.77</td>
<td>92.3</td>
</tr>
<tr>
<td>Identify</td>
<td>0.70</td>
<td>0.83</td>
<td>0.14</td>
<td>1.10</td>
<td>3.32</td>
<td>0.08</td>
<td>61.8</td>
</tr>
<tr>
<td>One-Back</td>
<td>-0.41</td>
<td>1.00</td>
<td>-0.18</td>
<td>1.10</td>
<td>0.48</td>
<td>0.48</td>
<td>0.22</td>
</tr>
<tr>
<td>Learn</td>
<td>-0.31</td>
<td>0.82</td>
<td>-0.48</td>
<td>0.85</td>
<td>0.40</td>
<td>0.53</td>
<td>0.21</td>
</tr>
<tr>
<td>CPAL</td>
<td>0.43</td>
<td>1.10</td>
<td>0.85</td>
<td>1.65</td>
<td>0.70</td>
<td>0.41</td>
<td>0.30</td>
</tr>
<tr>
<td>GML</td>
<td>-0.35</td>
<td>0.47</td>
<td>-0.30</td>
<td>0.55</td>
<td>0.08</td>
<td>0.78</td>
<td>0.09</td>
</tr>
<tr>
<td>ISL</td>
<td>-0.10</td>
<td>1.10</td>
<td>-0.18</td>
<td>1.17</td>
<td>0.04</td>
<td>0.85</td>
<td>0.06</td>
</tr>
<tr>
<td>ISL-R</td>
<td>-0.11</td>
<td>0.91</td>
<td>-0.29</td>
<td>0.90</td>
<td>0.38</td>
<td>0.54</td>
<td>0.20</td>
</tr>
</tbody>
</table>

In order to be consistent with Maruff and colleagues’ 2009 methodology, a series of t-tests were also performed in order to compute and evaluate differences in group performance without accounting for variance associated with premorbid intelligence, sleep problems, or time since transplant. The t-statistics were converted into Cohen’s $d$ by the following formula, suggested by Zakzanis (2002): $d = \frac{\text{Mean}_{	ext{Survivors}} - \text{Mean}_{	ext{Spouses}}}{\text{SD}_{	ext{Survivors}}}$ —. The %NOL statistic was then constructed using values proposed by Cohen (1988). Table 15 presents results of t-tests and associated %NOL.
Table 15

<table>
<thead>
<tr>
<th>Task</th>
<th>t</th>
<th>p</th>
<th>%NOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detect</td>
<td>-0.24</td>
<td>0.81</td>
<td>7.70</td>
</tr>
<tr>
<td>Identify</td>
<td>1.81</td>
<td>0.08</td>
<td>38.20</td>
</tr>
<tr>
<td>One-back</td>
<td>-0.71</td>
<td>0.48</td>
<td>14.70</td>
</tr>
<tr>
<td>Learn</td>
<td>0.64</td>
<td>0.52</td>
<td>14.70</td>
</tr>
<tr>
<td>CPAL</td>
<td>-0.88</td>
<td>0.39</td>
<td>21.30</td>
</tr>
<tr>
<td>GML</td>
<td>-0.29</td>
<td>0.78</td>
<td>7.70</td>
</tr>
<tr>
<td>ISL</td>
<td>0.23</td>
<td>0.82</td>
<td>7.70</td>
</tr>
<tr>
<td>ISL-R</td>
<td>0.64</td>
<td>0.53</td>
<td>14.70</td>
</tr>
</tbody>
</table>

As illustrated, %NOL ranged from 7.70 - 38.20. The %NOL represents the degree to which HCT survivor scores do not overlap with spousal partner scores. Identify, which exhibited a %NOL of 38.20, had the best discriminative ability of CogState tasks, as 38% of HCT survivors obtained scores on Identify unlike scores obtained by spousal partners. In accordance with study hypotheses, survivors performed worse than spouses on this task. CPAL, which had the next highest %NOL (21.30), discriminated survivors as performing better than spouses, contrary to study hypotheses.

**Criterion validation hypotheses.** As previously noted, the second series of study hypotheses were designed to address Adapted CogState Brief Battery criterion validity. Healthy control participant (e.g. spousal partner) performance on CogState tasks was expected to differ significantly from HCT survivor performance. Based on findings of Maruff and colleagues’ prior study (2009), as measured by the %NOL statistic, CogState tasks were expected to yield average %N-OLs of 50-60%.
a. The Detect task was hypothesized to yield a %NOL of 41-62% or greater. 
This hypothesis was not supported. The Detect task yielded a %NOL of 7.70.

b. The Identify task was hypothesized to yield a %NOL of 53-61% or greater. 
This hypothesis was not supported. The Identify task yielded a %NOL of 38.20.

c. The One-Back task was hypothesized to yield a %NOL of 55-60% or greater. 
This hypothesis was not supported. The One-Back task yielded a %NOL of 14.70.

d. The Learn task was hypothesized to yield a %NOL of 52-78% or greater. 
This hypothesis was not supported. The Learn task yielded a %NOL of 14.70.

e. Supplemental CogState tasks (e.g., CPAL, GML and ISLT) were 
   hypothesized to yield %NOLs of 41-78%. These hypotheses were not 
supported. CPAL yielded a %NOL of 21.30, GML a %NOL of 7.70, ISL a 
   %NOL of 7.70, and ISL-R a %NOL of 14.70.

Given that this set of hypotheses was not supported, a set of exploratory analyses were 
conducted in order to evaluate %NOL in traditional tasks. This was of interest, as the %NOL had 
not yet been explored in traditional tasks used to evaluate cognitive change in HCT survivors. 
See Table 16 for a summary of these results, alongside CogState tasks, organized according to 
domain.
Table 16

*T-test results and %NOL for HCT survivors and spousal partners, according to domain*

<table>
<thead>
<tr>
<th>Cognitive Domain</th>
<th>CogState</th>
<th>t</th>
<th>P</th>
<th>%NOL</th>
<th>Traditional</th>
<th>t</th>
<th>p</th>
<th>%NOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processing/ Psychomotor Speed</td>
<td>Detect</td>
<td>-0.24</td>
<td>0.81</td>
<td>7.70</td>
<td>GPB-D</td>
<td>-1.92</td>
<td>0.06</td>
<td>38.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GPB-ND</td>
<td>-0.52</td>
<td>0.60</td>
<td>14.70</td>
</tr>
<tr>
<td>Visual Attention</td>
<td>Identify</td>
<td>1.81</td>
<td>0.08</td>
<td>38.20</td>
<td>TMT-A</td>
<td>-0.99</td>
<td>0.33</td>
<td>21.30</td>
</tr>
<tr>
<td>Working Memory</td>
<td>One-Back</td>
<td>-0.71</td>
<td>0.48</td>
<td>14.70</td>
<td>SDMT</td>
<td>-1.86</td>
<td>0.07</td>
<td>38.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Span</td>
<td>-0.67</td>
<td>0.51</td>
<td>14.70</td>
</tr>
<tr>
<td>Visual Learning and Memory</td>
<td>Learn</td>
<td>0.64</td>
<td>0.52</td>
<td>14.70</td>
<td>RCFT</td>
<td>-0.42</td>
<td>0.68</td>
<td>7.70</td>
</tr>
<tr>
<td></td>
<td>CPAL</td>
<td>-0.88</td>
<td>0.39</td>
<td>21.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Executive Function</td>
<td>GML</td>
<td>-0.29</td>
<td>0.78</td>
<td>7.70</td>
<td>TMT-B</td>
<td>-2.21</td>
<td>0.03</td>
<td>43.00</td>
</tr>
<tr>
<td>Verbal Learning and Memory</td>
<td>ISL</td>
<td>0.23</td>
<td>0.82</td>
<td>7.70</td>
<td>HVLT</td>
<td>1.29</td>
<td>0.21</td>
<td>27.40</td>
</tr>
<tr>
<td></td>
<td>ISL-R</td>
<td>0.64</td>
<td>0.53</td>
<td>14.70</td>
<td>HVLT-R</td>
<td>1.35</td>
<td>0.18</td>
<td>27.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HVLT-RET</td>
<td>-0.04</td>
<td>0.97</td>
<td>0.00</td>
</tr>
</tbody>
</table>

As illustrated in Table 16, GPD-D, SDMT, and TMT-B evidenced modest discriminative ability, with %NOL 38.20 – 43.00. These scores indicate that 38.20 – 43.00% of HCT survivors obtained lower scores on the GPB-D, SDMT, and TMT-B than obtained by [healthy] spousal partners. HVLT and HVLT-R, which yielded a modest %NOL (27.40), discriminated spousal partners’ performance as worse than survivors, contrary to study hypotheses.

Of CogState tasks, Identify, which yielded the highest %NOL (38.2), discriminated spouses as performing better than survivors, as hypothesized. Contrary to study hypotheses, CPAL (%NOL = 21.3) discriminated survivors as performing better than spouses. Also contrary
to study hypotheses, ISL-R showed survivors as performing better than spouses (%NOL = 14.7), but this low %NOL suggests that the test did not discriminate well.

A final set of exploratory analyses was conducted in order to explore CogState criterion validity. A series of paired-samples t-tests were conducted to compare CogState task performance on HCT survivor/spousal partner on a pairwise basis, in order to evaluate whether differences emerged when survivors were compared against their own spousal partners, thereby roughly controlling for SES, pre-morbid IQ, and education. A significant difference emerged between matched HCT survivors and spousal partners on Identify, the CogState task of processing/psychomotor speed ($p > .05$), with survivors performing worse than spouses, in accordance with study hypotheses. See Table 17 for a summary of these paired-samples t-tests results.

Table 17

<table>
<thead>
<tr>
<th>Domain</th>
<th>CogState Task</th>
<th>$t$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual Attention</td>
<td>Detect</td>
<td>-0.22</td>
<td>0.83</td>
</tr>
<tr>
<td>Processing/Psychomotor Speed</td>
<td>Identify</td>
<td>2.27</td>
<td>0.04</td>
</tr>
<tr>
<td>Working Memory</td>
<td>One-Back</td>
<td>-0.67</td>
<td>0.51</td>
</tr>
<tr>
<td>Visual Learning and Memory</td>
<td>Learn</td>
<td>0.67</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>CPAL</td>
<td>-1.37</td>
<td>0.19</td>
</tr>
<tr>
<td>Executive Function</td>
<td>GML</td>
<td>-0.26</td>
<td>0.80</td>
</tr>
<tr>
<td>Verbal Learning and Memory</td>
<td>ISL</td>
<td>0.20</td>
<td>0.84</td>
</tr>
<tr>
<td>Verbal Learning and Memory</td>
<td>ISL-R</td>
<td>0.57</td>
<td>0.57</td>
</tr>
</tbody>
</table>

**Discussion**

This study was originally proposed to address problems associated with tracking cognitive change over time in HCT survivors. The literature suggests that HCT patients
experience cognitive deficits prior to transplant. These deficits appear to increase through approximately three months post-transplant, before returning to pre-transplant deficit levels at one year post-HCT. Cross-sectional studies have demonstrated impairment in HCT patients occurring prior to transplant (Andrykowski et al., 1992; Harder et al., 2005) and six months through four years post (Booth Jones et al., 2005; Harder et al., 2002). Prospective studies have supported these findings, collectively showing cognitive deficits in HCT patients prior to transplant, with further decline at six weeks (Friedman et al., 2009) and three months post-HCT transplant (Syrjala et al., 2004; Schultz-Kindermann et al., 2007; Scherwath et al., 2012), followed by a return to baseline at one-year post-transplant (Jacobs et al., 2007; Syrjala et al., 2009). To date, one longitudinal study has shown improvement beyond baseline levels of impairment on information-processing speed and executive function through five years post-transplant, whereas deficits in motor dexterity and verbal learning/retention ceased to improve beyond one year post-transplant (Syrjala et al., 2011).

A major limitation of existing prospective studies is the use of traditional neuropsychological tests. When used over short time intervals, these tests can be confounded by well-recognized practice effects (Lezak, 1983), which would mask impairment at later test points if participants improved on tests due to practice. Given the promise of CogState tasks to evaluate cognitive change over repeated testings with minimal practice effects (Collie et al., 2003; Falleti et al., 2006), the present study sought to evaluate the construct and criterion validity of the Adapted CogState Brief Battery in HCT survivors through replicating methodology of the validation study by Maruff and colleagues (2009). These authors found CogState tasks to be sensitive measures for detecting mild cognitive impairment associated with AIDS-dementia, mild traumatic brain injury, and schizophrenia. Mirroring the methods of Maruff and colleagues
the present investigation evaluated construct validity through analysis of associations between CogState tasks and traditional neuropsychological measures; criterion validity was evaluated through examination of whether CogState tasks could discriminate HCT survivors from their spousal partners. It was anticipated that if CogState tasks evidenced construct and criterion validity, use of the Adapted CogState Brief Battery, the Brief Battery expanded to include tasks of verbal learning and memory and executive function, could be useful for tracking cognitive changes in HCT survivors through repeated assessments over time.

Overall, study findings were mixed. With regard to construct validation, a number of tasks showed promise, but on the whole, correlations between CogState and traditional neuropsychological tasks measuring similar domains were lower than expected. In some cases, CogState tasks correlated with tasks in unexpected directions. With regard to criterion validity, CogState discriminability was also much lower than expected. The following sections will discuss findings, first reviewing construct validation results, followed by a review of criterion validation results. This discussion will be followed by conclusions about the overall validity of the Adapted CogState Brief Battery; implications, limitations, and strengths of the study; conclusions, and areas for future research.

As previously noted, two primary sets of hypotheses were proposed to address research questions associated with construct and criterion validity of the Adapted CogState Brief Battery.

1. **Construct Validation Hypotheses.** It was hypothesized that HCT survivors’ CogState task scores would correlate significantly with scores of traditional tests purported to measure similar neuropsychological domains.

   a. **Processing/psychomotor speed.** Performance on the CogState Detect task, a measure of processing/psychomotor speed, had been hypothesized to demonstrate the strongest
associations with tasks requiring simple visual attention, or psychomotor functions: GPB-D, GPB-ND, and TMT-A. This hypothesis was partially supported. Detect was most strongly associated with these tasks in predicted directions, although relationships with GPB-D and GPB-ND fell short of significance. It did evidence a statistically significant relationship with TMT-A, as predicted.

b. **Visual attention.** Performance on the Identify task, a measure of visual attention, had been hypothesized to demonstrate the strongest associations with tasks requiring visual processing and divided attention: TMT-A, TMT-B, and SDMT. This hypothesis was also partially supported. Identify was strongly negatively correlated with TMT-A as predicted, although this relationship only approached significance. It was also strongly correlated with TMT-B in the predicted direction, but this relationship also failed to reach significance.

c. **Working memory.** Performance on the One-Back task, a measure of working memory, was hypothesized to demonstrate the strongest associations with tasks requiring visual scanning and working memory: SDMT, Span, and TMT-B. This hypothesis was not supported, as the task evidenced a weak and statistically non-significant relationship with all of these tasks. Notably, its relationship to traditional tasks of verbal learning (i.e. HVLT) and visual attention (i.e. TMT-A) were strong.

d. **Visual learning and memory.** Performance on the Learn and CPAL tasks, measures of continuous visual recognition learning, was hypothesized to demonstrate the strongest associations with a task requiring visual learning and memory: RCFT-R. This hypothesis was partially supported. Learn was only moderately correlated with RCFT-R. Rather, it showed a strong relationship to a task of psychomotor speed (e.g., GPB) and moderate
relationships with working memory tasks (e.g., SDMT and Span). CPAL, on the other hand, did correlate strongly and negatively as predicted with RCFT, although the relationship fell short of significance. Curiously, it evidenced a negative relationship with TMT-B, a traditional measure of executive function and visual attention.

e. **Executive function.** Performance on the GML task, a measure of executive functioning, was hypothesized to demonstrate the strongest associations with traditional tests requiring visual processing and mental flexibility: SDMT and TMT-B. This hypothesis was not supported: relationships between GML and each of these tasks were weak. Although its relationship with TMT-B was negative as predicted, its relationship with SDMT was positive, contrary to study hypotheses.

f. **Verbal learning and memory.** Performance on the ISL, a measure of verbal learning and memory, was expected to demonstrate the strongest associations with HVLT, a task requiring verbal learning and memory. This hypothesis was supported; these tasks showed strong inter-correlations.

Overall, analysis of Pearson product-moment correlations provided partial support for construct validation hypotheses: Detect, Identify, CPAL, and ISL showed at least some strong relationships with tasks in expected directions. Exploratory analyses of CogState convergence and divergence illustrated expected patterns among CogState tasks as well as some traditional measures.

**Effect Size Analyses**

Effect size analyses, including effect size computation, plotting, and graphing with confidence intervals, showed more overlap among CogState tests than prior validation work indicated. In the original CogState validation study, Maruff and colleagues’ (2009) effect size
analysis revealed very little overlap, suggesting that CogState tasks measure distinct, non-overlapping constructs, according to Cumming and Finch’s recommendations for analyzing confidence interval overlap (2005). However, in the current study when statistically significant correlations between CogState and traditional tasks were transformed into Cohen’s $d$ effect sizes and graphed, results revealed overlap across three CogState tasks’ (i.e. ISL-R, One-Back, and Detect) effect sizes with traditional measures, HVLT and TMT-A. HVLT effect sizes with ISL-R and One-Back exhibited confidence interval overlap; TMT-A effect sizes with Detect and One-Back likewise exhibited overlap. Additional analyses, within which Pearson Correlation Coefficients $\geq .30$ were transformed into Cohen’s $d$ effect sizes and graphed, revealed overlap across all seven tasks. RCFT effect sizes with Identify and CPAL exhibited confidence interval overlap; HVLT effect sizes with One-Back and ISL exhibited overlap; HVLT-RET with ISL, ISL-R, and One-Back; SDMT with CPAL, Learn, and ISL-R; TMT-A with Identify, Detect, and One-Back; GPB-D with CPAL and Identify; GPB-ND with ISL-R and CPAL.

These findings were not surprising from a broader neuropsychological perspective, in that neuropsychological tasks typically do not measure distinct constructs, and overlap across tasks is expected (Lezak, 1983). However, it remains curious that present study findings were so disparate from Maruff and colleagues’ original 2009 validation study, which revealed less overlap: CogState tasks showed distinct patterns of overlap with traditional tasks but not each other. Just why the patterns of overlap shown in the present study did not mirror Maruff and colleagues’ remains unclear. One possible explanation for this discrepancy is sample size. In contrast to the original validation work by these authors in 120 participants (50 schizophrenic, 50 mild traumatic brain injury, and 20 AIDS-dementia), this study’s 40 participants provided sufficient power to detect large effects sizes only. For a pilot study including face-to-face testing
of a rare medical population, the investigator had deemed this a priori to be sufficient power to contribute to the field. However, that appears in fact to have been not large enough to detect moderate or smaller correlational relationships.

Another possibility is that CogState tasks were simply not sensitive enough to detect subtle patterns of impairment associated with HCT. Patterns of deficits have been shown in comparison with population norms, but, in general, changes experienced by HCT survivors are quite subtle, and researchers have used varying criteria to classify impairment: Booth-Jones et al. (2005) and Syrjala et al. (2011) used impairment criterion of $z \leq 1.0$; Scherwath et al. (2012) and Schultz-Kindermann et al. (2007) used impairment criterion of $z \leq 1.4$; Andrykowski et al. (1994), Harder et al. (2005), Jacobs et al. used $z \leq 1.5$; Harder et al. (2002) and Meyers et al. (1994) used $z \leq 2.0$. This would not be surprising, given that Hammers et al. (2011) also found CogState tasks to have less than optimal discriminability. As previously noted, these authors found CogState tasks were unable to discern mild cognitive impairment subjects from healthy controls.

Traditional neuropsychological tasks used in this study demonstrated more cognitive deficits found by Syrjala et al. (2011), who also evaluated HCT survivors through five years post-HCT. Classifying impairment in $z$-scores $\leq 1.0$, Syrjala and colleagues identified impairment in approximately 7% of survivors on Digit Symbol at five years post-HCT, 12% on HVLT, 33% on GPB-D, 26% on GPB-ND, 8% on TMT-A, and 8% on TMT-B. Using this same cutoff criterion in the present study ($z \leq 1.0$), 20% of HCT survivors met criteria for impairment on SDMT, 25% on HVLT, 70% on GPB-D, 60% on GPB-ND, 25% on TMT-A, and 35% on TMT-B. Although it is unclear why impairment was higher in the present study, it is possible that this difference was due to sampling. The present study recruited entirely through online and
in-person support groups. Consequently, those interested in participating might have been struggling more than those recruited in a hospital setting, such as within Syrjala’s sample (2011).

Using Strauss and Spreen’s recommended cutoff criterion of $z \leq 1.5$ (2006), 5% of survivors in the present study showed impairment on SDMT, 15% on HVLT, 40% on GPB-D, and 35% on GPB-ND. Thus, it appears that, in the context of the small literature investigating HCT patients through five years post-HCT, the present study sample has more cognitive deficits than what has been found previously. Our findings do support those of Syrjala et al., which demonstrated long-lasting impairment in memory and processing/psychomotor speed (2011).

This degree of impairment could have significant implications for those suffering HCT-associated declines. In terms of quality of life, subtle changes such those illustrated by participants in this study are likely quite frustrating. The general cancer literature suggests that cognitive changes associated with chemotherapy have a large impact on quality of life (Ahles & Saykin, 2007). Patients experiencing cognitive changes associated with chemotherapy report disturbances in their activities of daily living (Tannock et al., 2004), particularly during times of multi-tasking.

This is not surprising, given the functions impacted by chemotherapy, such as that used in HCT. Psychomotor speed, the area most heavily effected in the present sample, measures the “speed and accuracy with which a person can perform simple motor tasks and manipulate objects” (Ahles & Saykin, 2007). Thus, the degree to which deficits might impact an individual are broad and could range from experiencing difficulty picking up small objects or buttoning a shirt to potentially performing on the job. Processing speed, which according to Ahles and Saykin (2007) refers to “the speed or efficiency with which information can be used in completing a task,” might also impact a person’s ability to perform in a variety of settings.
broadly, from work environments to functioning at home, and a person experiencing a decline in processing speed might sense a shift in overall efficiency. Finally, deficits in verbal memory, or “memory for words and narrative material presented verbally or in writing,” (Ahles & Saykin, 2007), shown by participants in this sample could have a large impact on an individual’s ability to function across many domains, at work or recreationally, depending on the degree to which verbal expression is required in these settings.

Although literature documenting specific difficulties experienced by individuals with mild deficits in the aforementioned domains is still under development, it is likely that the impact of such deficits are broad, particularly in work environments which require attention to detail and quick response times. Deficits in processing/psychomotor could have an especially broad impact on one’s occupational functioning. Subtle declines in this domain might, for example, manifest in a slower work production, including responses to assigned tasks, communication abilities, and overall efficiency. Declines in verbal fluency/memory, which manifests as word-retrieval difficulty, could also impact an individual’s effectiveness in communication, in both oral and written formats. Difficulties of this nature could jeopardize an individual’s ability to work effectively and, in some cases, make it impossible to perform tasks that were once easy to complete. Anecdotally, a number of HCT survivors who participated in the present study reported difficulties at work. For example, one reported that she was completely unable to function in her position as medical coder after transplant and had been forced to resign from the position entirely.

Cysique et al. (2006) and Overton et al. (2011) both found CogState Detect and Identify tasks to have strong sensitivity with regard to identifying deficits in psychomotor processing and reaction speed performance. Given that our sample had such a large degree of impairment in
these domains, it is puzzling that deficits were not captured by CogState tasks, even when impairment criterion was set at $z \leq 1.0$. These results further suggest that CogState lacks sensitivity needed to detect HCT-associated cognitive impairment.

Regardless, it remains unclear as to why this study evidenced a larger degree of overlap between CogState tasks and traditional measures than that shown in original validation work by Maruff and colleagues (2007). One possibility is that CogState normative data that might affect confidence intervals was still in development at the time of Maruff et al.’s original study. If CogState tasks norms changed significantly, this might impact the degree to which task effect sizes and confidence intervals overlap when graphed. To date, Adapted CogState Brief Battery normative data remains under construction (personal communication, Adrian Schembri, see Appendix M).

It also remains unclear as to why other correlations between CogState and traditional tasks measuring similar traits were weaker and, in some cases, correlated in opposite directions than predicted; even when criteria were expanded, correlations were much weaker than predicted. One possible explanation may be that CogState tasks evaluated in this study are different from those this investigator expected a priori they would measure.

With regard to Identify, the CogState task of visual attention, it is unclear as to why correlations were high with some measures of visual attention (e.g. TMT-A) but not with others (e.g., Span and SDMT). The strong inter-correlation between Identify and Detect casts doubt on the utility of Identify as a pure measure of visual attention.

In contrast, One-Back, a CogState measure of working memory, exhibited a weaker-than-predicted relationship with traditional measures of working memory (e.g., SDMT, Span, and
TMT-A), but it did show a strong relationship to a traditional verbal learning and memory task (e.g., HVLT) as expected.

Learn, a CogState task of visual learning and memory, which was expected to correlate most strongly with a traditional task of visual learning and memory (RCFT), showed strong relationships with traditional measures of working memory (e.g. SDMT and Span), which also draw on visual learning. This indicates that Learn could be used broadly as an indicator of visual learning or memory in HCT survivors.

CPAL, purported to be a measure of visual learning and memory, did exhibit a strong relationship to RCFT as expected. Its relationship with TMT-B was also strong, but in the opposite direction predicted.

2. Criterion Validation Hypotheses. With regard to criterion validation, hypotheses were unsupported overall. Findings suggested that CogState discriminative ability is much lower than expected and that the Adapted CogState Brief Battery does not possess discriminative ability necessary to differentiate clinical HCT candidates and survivors from healthy controls (e.g. criterion validity).

Interestingly, when well-established traditional neuropsychological measures were subjected to %NOL analyses, resulting criterion validity was much lower than that reported in prior CogState validation work (Maruff et al., 2009). Thus, the issue is not necessarily with CogState tasks, but with difficulties discriminating between HCT survivors up to 5 years post-HCT from healthy controls. It is worth noting that is the first study of its kind to evaluate validity of any form of neuropsychological measurement in HCT survivors. Of CogState tasks, Identify demonstrated the highest %NOL statistic (38.20), discriminating better performance in spousal partners. This discriminability is in the same range or better than traditional tasks such as TMT-
A, which demonstrated a %NOL of 21.30. As such, Identify may be seen as possessing discriminative ability at least as good as traditional neuropsychological assessment. CPAL, which also demonstrated a %NOL of 21.3, discriminated HCT survivors as performing better than spouses.

Exploratory analyses, and in particular, construction of %NOL for all tasks used in this study suggested that traditional neuropsychological measures possess more sensitivity than CogState in discriminating HCT survivor scores from those of healthy control subjects (e.g. spousal partners). Although sensitivity of these traditional tasks was greater than that evidenced by CogState, the traditional, well-established tasks reflected only modest discriminative ability. As noted in study results, TMT-B, a traditional measure of executive function, possessed the strongest discriminative ability across all tests administered in this study, exhibiting a %NOL of 43.00, indicating that 43% of HCT survivors scored differently than spousal partners on this test. This finding is consistent with other literature demonstrating TMT-B discriminability (Bastug et al., 2013). However, it is important to note that this well-established measure discriminates less than ideally, according to standards defined by Zakzanis (2001), who suggests that a %NOL of 93 is indicative of a good clinical marker. This suggests measurement through the use of well-established neuropsychological measures likely also lack the criterion validity (i.e. discriminability) necessary to distinguish subtle cognitive changes experienced by HCT survivors from cognitive change experienced by the general population. HCT related-cognitive deficits are difficult to evaluate, and at present there is a need for highly sensitive measures that can be used over a multiple assessments. Prior research has indicated that CogState tasks are useful in tracking cognitive change over multiple assessments with minimal practice effects (Falletti et al., 2006; Maruff et al., 2009). However, in light of the lower than expected
relationships with traditional cognitive measures, it appears that they may lack sensitivity necessary to detect HCT-associated cognitive decline in patients up to 5 years after transplant in comparison to their spouses roughly matched for SES, education, and emotional toll of transplant.

**Implications**

Taken together, CogState did not perform as predicted and exhibited limited sensitivity to for detecting cognitive impairment long after HCT. However, given its previously established temporal stability, the measure may have some utility for tracking cognitive changes prior to transplant. To date, only one study (Syrjala et al., 2011) has documented cognitive recovery in HCT survivors through five years post HCT; research investigating HCT-associated cognitive function beyond this time point is limited. The most profound impairment in HCT survivors has been shown as occurring within the first 80 days (Schultz-Kindermann et al., 2007; Scherwath et al., 2012; Syrjala et al., 2011), after which functioning appears to return to pre-transplant levels (Harder et al., 2004; Syrjala et al., 2004, 2012) in most neuropsychological domains, although deficits have still been exhibited in survivors one year post-transplant. Although cognitive deficits are well documented in 15-46% of HCT survivors at one year post-transplant (Harder et al., 2002, 2004; Syrjala et al., 2004; Meadows et al., 2013), prospective studies have revealed that impairment at this time point has largely recovered to pre-transplant levels and is quite mild (Jacobs et al., 2007; Syrjala et al., 2004, 2011). Further research investigating the utility of the Adapted CogState Brief Battery at early time points is warranted, since the greatest degree of declines have been shown early on in the transplant trajectory (Beglinger et al., 2006; Syrjala et al., 2004). Thus, it is possible that declined functioning occurring early on may be more discernible by CogState tasks.
Furthermore, since cognitive function appears to recover to baseline levels within one year (Jacobs et al., 2007, Syrjala et al, 2004, 2011), it is also likely that deficits exhibited in the were too subtle for CogState to detect, particularly since HCT survivors’ cognitive functioning appears to continue recovering through five years post-HCT (Syrjala et al., 2011). It is, however, also important to note that these reported recovery trends might be contaminated by practice effects. Repeated administration of CogState could clarify this; possible future investigation could analyze data using a within-subjects design as did Friedman et al., (2009), whereby subject baselines could serve as a point for evaluating change.

Given CogState’s capacity for remote computer administration, there is a potential for patients to periodically complete key CogState tasks following their discharge from the hospital, in order to for doctors to track cognitive changes over time, since declines in cognitive function could indicate poor medical outcome (e.g., delirium associated with infection).

Use of the Adapted CogState Battery as a screening tool in general for evaluating patients prospectively over the course of transplant is another alternative. Recently, joint recommendations of the European Group for Blood and Marrow Transplantation, the Center for International Blood and Marrow Transplant Research, and the American Society of Blood and Marrow Transplantation called for a for a high vigilance for psychological symptoms, including monitoring of cognitive symptoms (Majhail et al., 2012). More specifically, the group recommended that HCT recipients undergo assessment of cognitive function at yearly intervals. Given difficulty associated with traditional neuropsychological measurement (e.g. administration via highly-trained psychometrician, potential for practice effects if repeated assessment within 6 months, timeliness), use of select CogState tasks may offer one way of meeting this recommendation. HCT candidates and survivors could effectively be monitored through the
duration of transplant, which would allow doctors to closely monitor cognitive change over time and respond accordingly. Use of CogState tasks could provide valuable information for determining possible courses of intervention (i.e. cognitive rehabilitation/retraining or pharmaceutical) to address cognitive declines if warranted. Repeated assessment with CogState tasks could also be useful in tracking changes following such interventions.

**Limitations**

The current study had several limitations that must be noted. Unlike other investigations of cognitive functioning in HCT, this study was conducted entirely outside of the hospital setting. Participants responded to ads posted on HCT support websites or information provided during support group meetings. Thus, our study recruited socially connected HCT survivors and their partners. To date, no published research has evaluated the use of social networking/support group utilization across HCT survivors, although preliminary findings of a large, online-dataset investigating help-seeking behavior in HCT survivors (Gourley et al., 2011) suggest that a considerable number of HCT survivors feel a need for supportive services beyond three years post-transplant. Given that recruitment targeted online and community-based support groups, it is possible that participants in this study were among a subset up survivors experiencing greater post-transplant struggle, thus feeling the need to reach out for support through websites and community services. If this is the case, the present study sample might be skewed toward those experiencing high levels of distress long after transplant and may not be generalizable to the HCT population at large. However, it is also worth noting that all other neurocognitive studies of HCT survivors have been hospital-based and not representative of the HCT population as a whole.
The present study sample may also have been confounded by gender sampling differences. Although no studies have identified an association between increased impairment and gender in HCT survivors, 60% of HCT survivors in the present study were females. These differences in male and female participation limit the generalizability of study findings.

The small sample size of this study is also a limitation. The power analysis (Cohen, 1988) undertaken during design of this study suggested that an n = 26 would allow for detection of a large effect size. A large effect size had been expected; however, given that other researchers who had evaluated HCT survivors against matched controls detected few differences between groups (Syrjala et al., 2011), it is not surprising that few differences between participant groups in this study emerged. A larger sample should have been recruited in order to determine whether subtle group differences emerged (i.e. avoid a Type II error). From a control group standpoint, spousal partners were effective in controlling for age, premorbid IQ, and socioeconomic and psychological variables: the only difference that emerged between groups was poorer sleep for HCT survivors, although variance associated with this difference did not account for performance differences between HCT survivors and spousal partners.

**Strengths**

This study has a number of strengths. To date, this is the first to use HCT survivors and spousal partners as matched-controls. Given that no significant differences between groups emerged across socioeconomic and psychosocial measures, including premorbid IQ, our results suggest that spousal partners function well as controls for HCT survivors. These findings suggest that use of spousal partners as matched controls may be advantageous in future investigations of cognitive and psycho-social functioning in HCT survivors across the transplant trajectory. However, it is also worth noting that using a criterion cutoff of $z \leq 1.5$, 30% of spousal partners
exhibited impaired performance on GPB-ND, a traditional task of processing/psychomotor speed, and 20% exhibited impairment on RCFT-R, a traditional task of visual learning and memory. Given this high degree of impairment, one could speculate that stress of HCT takes a significant toll of spouses’ cognitive function, and, thus, spousal partners should be used with caution. In particular, spousal partners performed worse on Identify than survivors. Although the %NOL (21.3) was still lower than needed for this task to be a good clinical marker, it may be an indicator of the cognitive impact of stress in HCT spouses. In order to further clarify cognitive deficits associated with stress in spousal partners, future studies should investigate spouse functioning alongside peer-nominated control subjects.

Further, although validity of the Adapted CogState Brief Battery was not as strong as predicted, this study represents the first of its kind to evaluate the validity of neuropsychological test use in HCT survivors. Given potential difficulties associated with detecting subtle changes associated with HCT, future studies evaluating the validity of both computerized and traditional neuropsychological tests are warranted, especially early post-transplant.

Finally, unique recruitment of HCT survivors outside of the hospital setting is also a strength of this study, in that it provides additional information about cognitive functioning of HCT survivors and their spousal partners outside of the hospital settings. The present study findings demonstrated higher degrees of cognitive impairment than what was found in previous work where participants were recruited in a hospital setting (Syrrjala et al., 2011). These findings are important in that it is possible that post-transplant HCT survivors are in need of more support than is available outside of the hospital setting. Although long-term psychological sequelae are recognized in both HCT survivors and their spousal partners, mechanisms are not yet in place for effectively delivering services to individuals in need outside of the hospital setting.
Conclusion

In conclusion, although CogState performed differently than expected, various tasks, particularly Identify, appear to have a certain degree of utility and may be useful for tracking cognitive change over time in HCT candidates and survivors. The use of CogState within this arena could provide a feasible means for doctors to assess patients on a weekly or monthly basis. Given the simple CogState interface, patients could complete repeated testings independent of a psychometrist in order to provide doctors valuable information about their cognitive status as they progress through the transplant trajectory.

Further, since some CogState tasks did correlate with well-established traditional tests that are seen to be valid indicators of cognitive deficit on cross-sectional testing, there is reason to further evaluate the utility of this measure, particularly since traditional tests are unable to be used repeatedly.

Future Directions

Although present study findings did not show the predicted strong associations between Adapted CogState Brief Battery tasks and traditional neuropsychological measures, relationships that approached significance suggest that the battery has some utility for use with HCT survivors. Further, given its previously established test-retest reliability (Falleti et al., 2006; Collie et al., 2003), investigation of CogState tasks across the transplant trajectory could be helpful in further evaluating the utility of this tool for tracking cognitive decline and recovery in HCT survivors.

Repeated measurement at key time points during the transplant trajectory using various tasks from the Adapted CogState Brief Battery would be of particular value, with a goal of evaluating patterns of cognitive change. Examination of CogState repeated measurement could
be helpful in confirming or disconfirming patterns of cognitive recovery at one year post-transplant shown in prior prospective studies (Syrjala et al., 2004; 2011; Jacobs et al., 2007), while sidestepping the potential confound of practice effects of traditional measures.
References


Gourley, B., Hoodin, F., Carey, J., Alschuler, K., Hart, L., & Sprague, C. (2011). How can we move forward if we don’t know where we are in providing mental health services (MHS) to Hematopoietic Stem Cell Transplant (HCT) patients? [Abstract] *Annals of Behavioral Medicine, 41*, s156.


McQuellon R., Russell, G., Rambo, T., Craven, B., Radford, J., Perry, J., Cruz, J., & Hurd, D. (1998). Quality of life and psychological distress of bone marrow transplant...
recipients: the “time trajectory” to recovery over the first year. *Bone Marrow Transplant*, 21, 477-486.


Patient-rated emotional and physical functioning among hematologic cancer patients
during hospitalization for stem-cell transplantation. *Bone Marrow Transplantation*, 35,
307–314.

Psychiatric morbidity and impact on hospital length of stay among hematologic cancer
patients receiving stem-cell transplantation. *Journal of Clinical Oncology*, 20, 1907-
1917.

Clinical factors associated with fatigue in haematologic cancer patients receiving stem-

San Antonio, TX: Author.

Assessment of Neuropsychological Status (RBANS): Preliminary clinical validity.


Riva, D., Nichelli, F., & Devoti, M. (2000). Developmental aspects of verbal fluency and
Rizzo, J., Wingard, J., Tichelli, A., Lee, S., Van Lint, M., Burns, L., Davies, S.,
preventative practices for long-term survivors after hematopoietic cell transplantation:
joint recommendations of the European Group for Blood and Marrow Transplantation,
Center for International Blood and Marrow Transplant Research, and the American
Society for Blood and Marrow Transplantation (EBMT/CIBMTR/ASBMT). Bone
Marrow Transplantation, 37, 249-261.

Ross, T. (2003). The reliability of cluster and switch scores for the Controlled Oral Word

Ruff, R., Evans, R., & White, R. (1986). Automatic detection vs. controlled search: a paper and


Scherwath, A., Schimer, L., Kruse, M., Ernst, G., Eder, M., Dinkel, A., Kunze, S., Balck, F.,
Bornhauser, M., Ehninger, G., Dolan, K., Gramatzki, M., Hans-Jochem, K., HeuBner, P.,
(2012). Cognitive functioning in allogeneic hematopoetic stem cell transplantation
recipients and its medical correlates: a prospective multicenter study. Psycho-Oncology,
DOI: 10.1002.pon.3159.

role of biomedical and psychosocial factors for the prediction of pain and distress in
patients undergoing high-dose therapy and BMT/PBSCT. Bone Marrow Transplantation,
29, 341-351.


Zakzanis, K. (2001). Statistics to tell the truth, the whole truth, and nothing but the truth: formulae, illustrative numerical examples, and heuristic interpretation of effect size
analyses for neuropsychological researchers. *Archives of Clinical Psychology, 16*, 653-667.


Appendices
Appendix A: CogState Validation Study Eligibility Checklist

Inclusion criteria

Yes__ No__  Speak, read, and understand English

Yes__ No__  Are 18 years of age or older

Yes__ No__  Are mentally competent to provide informed consent

Yes__ No__  Have undergone or have a spousal or domestic partner who has also undergone a bone marrow or stem cell transplant within the past five years

Yes__ No__  Have a spousal or domestic partner also interested in participating

Exclusion criteria

Yes__ No__  Have an existing neurological disorder

Yes__ No__  Prior head injury

Yes__ No__  Current substance dependence

Yes__ No__  History of stroke, epilepsy, or other CNS pathology requiring radiation or surgery

Yes__ No__  History of brain tumor(s)
Appendix B: Consenting Script

To be read to potential participants:

We appreciate your interest in our study. If you do agree to participate, we will be examining your ability to complete a series of tasks. Some tasks will be performed while interacting with a test administrator, others using a computer.

The entire task set (including both computerized and non-computerized tasks) will take approximately one hour for each of you to complete.

There are no anticipated risks associated with this study. It is our hope that information gained will aid in our understanding of the experiences of bone marrow transplant patients.

Participating within this study is entirely voluntary.
Appendix C: Informed Consent- Survivors

Project Title: Validation of the Adapted CogState Brief Battery in hematopoietic cell transplant survivors

Investigator: Bethany Gourley, Eastern Michigan University

Co-Investigator: Flora Hoodin, PhD, Eastern Michigan University

Purpose of the Study: The purpose of this research study is to examine the usefulness of a set of brief, computerized neuropsychological tests in hematopoietic cell transplant (HCT) survivors. This information may help researchers better understand the neuropsychological impact of HCT and possibly also design supportive services for patients during and after transplantation.

Procedure: Taking part in this study will involve completing various tasks. Some of the tasks will be completed with a test administrator, while others will involve the use of a computer. In total, the tasks will take approximately one hour to complete and can be completed at the Eastern Michigan University Psychology Clinic, your local library, or within your home.

If you decide to participate in this study, you will:
   a. Answer a brief paper and pencil questionnaire about your transplant and personal history. It also includes questions about depression, anxiety, substance use and sleep patterns
   b. Spend 25 minutes doing computerized neuropsychological tests
   c. Spend 35 minutes doing neuropsychological tests with a test administrator

In total, approximately one hour will be spent completing tasks, on one occasion only.

In order to participate in this study, you must:
   • Be at least 18 years of age
   • Speak English
   • Have received a stem cell transplant within the last five years
   • Have a spousal or domestic partner also interested in participating
   • Have no history a neurological disorder
   • Have no history of head injury
   • Have no history of stroke
   • Have no history of epilepsy
   • Have no history of brain tumor or other brain condition requiring radiation or surgery
   • Have no present drug or alcohol dependence

Confidentiality: At no time will your name be associated with your test responses. Only a code number will identify your transcribed and computerized responses. The results will be stored separately from the consent form, which includes your name and any other identifying information.

All related materials will be kept in locked file cabinets in the researcher’s office. Your non-identifiable electronic data will be stored within a password-protected remote database.
**Expected Risks:** There are minimal foreseeable risks to participating in this study. For some, testing may be mildly distressing, although no harm is anticipated. You may choose to stop participating at any time and may notify the research assistant if you are having trouble with a test. As with any research study, there may be additional risks that are unknown or unexpected.

**Expected Benefits:** While you will not directly benefit from participating in this study, it is our hope that information gained will help future transplant patients and survivors.

**Voluntary Participation:** Participation in this study is voluntary. You may choose not to participate.

If you do decide to participate, you may change your mind at any time and withdraw from the study without negative consequences.

**Compensation:** Unfortunately, we are not able to compensate you for your participation, but we do value your input and information.

**Use of Research Results:** Results will be presented when all data is collected. No names or individually identifying information will be revealed. Results may be presented at research meetings and conferences, in scientific publications, or as part of a doctoral thesis being conducted by the principal investigator.

**Future Questions:** If you have any questions concerning your participation in this study now or in the future, you can contact the principal investigator, Bethany Gourley, via phone (734) 487-4987 or e-mail (bgourley@emich.edu) or the project supervisor, Flora Hoodin, PhD, via phone (734) 487-1155 or email (fhoodin@emich.edu).

This research protocol and informed consent document has been reviewed and approved by the Eastern Michigan University Human Subjects Review Committee for use from _____________ to _____________ (date). If you have questions about the approval process, please contact Dr. Deb de Laski-Smith (734.487.0042, Interim Dean of the Graduate School and Administrative Co-Chair of UHSCR, human.subjects@emich.edu)

**Consent to Participate:** I have read or had read to me all of the above information about this research study, including the research procedures, possible risks, side effects, and the likelihood of any benefit to me. The content and meaning of this information has been explained and I understand. All my questions, at this time, have been answered. I hereby consent and do voluntarily offer to follow the study requirements and take part in the study.
<table>
<thead>
<tr>
<th>PRINT NAME:</th>
<th>Signatures:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant (your signature)</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td></td>
</tr>
<tr>
<td>Investigator or Specified Designee</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td></td>
</tr>
</tbody>
</table>
Appendix D: Informed Consent-Partners

Project Title: Validation of the CogState Brief Battery in Hematopoietic Cell Transplant Survivors

Investigator: Bethany Gourley, Eastern Michigan University

Co-Investigator: Flora Hoodin, PhD, Eastern Michigan University

Purpose of the Study: The purpose of this research study is to examine the use of a brief computerized testing battery in hematopoietic cell transplant (HCT) survivors and their spouses or domestic partners. This information may help us design supportive services to help patients during and after transplantation as well as understand the neuropsychological impact of HCT.

Procedure: A research assistant will explain the study to you, answer any questions you may have, and witness your signature to this consent form.

In order to participate in this study, you must:
- Be the domestic partner or spouse of a stem cell transplant survivor, who has undergone transplantation within the last five years
- Speak English
- Be at least 18 years of age
- Not have an existing neurological disorder
- Not have previously suffered a head injury
- Not have a history of stroke, epilepsy, or other CNS pathology requiring radiation or surgery
- Not currently have alcohol or drug dependence
- Not have a history of brain tumor(s)

Testing: Taking part in this study will involve completing various tasks. Some of the tasks will be completed with a test administrator, while others will involve the use of a computer. In total, the tasks will take approximately one hour to complete and can be completed within your home or at the Eastern Michigan University Psychology Clinic.

Confidentiality: Only a code number will identify your transcribed and computerized responses. The results will be stored separately from the consent form, which includes your name and any other identifying information. At no time will your name be associated with your responses.

All related materials will be kept in locked file cabinets in the researcher’s office. Your non-identifiable electronic data will be stored within a password-protected remote database.

Expected Risks: There are minimal risks to participating in this study. For some, testing may be mildly distressing, although no harm is anticipated. You may choose to stop participating at any
time and may notify the research assistant if you are having trouble with a test. As with any research study, there may be additional risks that are unknown or unexpected.

**Expected Benefits:** While you will not directly benefit from participating in this study, it is our hope that information gained will help future transplant patients and survivors.

**Voluntary Participation:** Participation in this study is voluntary. You may choose not to participate.

If you do decide to participate, you may change your mind at any time and withdraw from the study without negative consequences.

**Compensation:** Unfortunately, we are not able to compensate you for your participation, but we do value your input and information.

**Use of Research Results:** Results will be presented when all data is collected. No names or individually identifying information will be revealed. Results may be presented at research meetings and conferences, in scientific publications, or as part of a doctoral thesis being conducted by the principal investigator.

**Future Questions:** If you have any questions concerning your participation in this study now or in the future, you can contact the principal investigator, Bethany Gourley, via phone (734) 487-4987 or e-mail (bgourley@emich.edu).

This research protocol and informed consent document has been reviewed and approved by the Eastern Michigan University Human Subjects Review Committee for use from ____________ to ____________ (date). If you have questions about the approval process, please contact Dr. Deb de Laski-Smith (734.487.0042, Interim Dean of the Graduate School and Administrative Co-Chair of UHSCR, human.subjects@emich.edu)

**Consent to Participate:** I have read or had read to me all of the above information about this research study, including the research procedures, possible risks, side effects, and the likelihood of any benefit to me. The content and meaning of this information has been explained and I understand. All my questions, at this time, have been answered. I hereby consent and do voluntarily offer to follow the study requirements and take part in the study.

PRINT NAME:
Signatures:

___________________________________________________  Date
Participant (your signature)  

___________________________________________________  Date
Investigator or Specified Designee  


Appendix E: Demographics Questionnaire for HCT Survivors

1. How old are you?

2. Are you
   a. Male
   b. Female

3. What is your ethnicity?
   a. Not Hispanic or Latino
   b. Hispanic or Latino

4. Some people identify themselves as belonging to one or more racial groups. Please indicate which of the following groups you belong to. Please check all that apply.
   a. White or Caucasian
   b. Black or African-American
   c. Hispanic or Latino
   d. American Native
   e. Alaskan Native
   f. Asian
   g. Pacific Islander
   h. Middle Eastern
   i. Other ______

5. How many years of education have you completed? (Completing high school or its equivalent = 12 years)

6. What is the highest level of education you have obtained?
   a. None
   b. High School Diploma
   c. GED
   d. Associates
   e. Bachelors
   f. Masters
   g. Doctorate
   h. Other (please specify)

7. What is your current marital status?
   a. Married
   b. Single
   c. Divorced
   d. Remarried
   e. Widowed
   f.Separated
g. Living with partner (same sex)
h. Living with partner (opposite sex)

8. What is your current employment status? (Please check one)
   a. Full Time (>35 hrs/wk)
   b. Part Time (Regular hours)
   c. Part Time (Irregular hours)
   d. Unemployed, full time student
   e. Unemployed, part time student
   f. Unemployed due to health issues related to transplant
   g. Retired for transplant related issues
   h. Retired
   i. On disability for transplant related issues
   j. On disability for other reasons
   k. Military service

9. What is the current economic status of your household? (Please check one)
   a. We have barely enough to get by
   b. We have enough to get by, but no more
   c. We are solidly middle class
   d. We have plenty of “extras”
   e. We have plenty of “luxuries”
   f. Don’t know/unsure/prefer not to say

10. What is your annual household income? (Please check one)
    a. >$150,000
    b. $100,000-$149,000
    c. $75,000-$99,000
    d. $50,000-$74,000
    e. $25,000-$49,000
    f. $10,000-$24,000
    g. <$9,000
    h. Don’t know, or prefer not to say

11. What was the date of your transplant? __________

12. What was/were your diagnosis/es at the time of your transplant? Please check all that apply.
    a. Acute lymphocytic leukemia (ALL)
    b. Acute myelogenous leukemia (AML)
    c. Acute osteomyelofibrosis
    d. Chronic lymphocytic leukemia (CLL)
    e. Chronic myeloid leukemia (CML)
    f. Fagoni’s anemia
    g. Hodgkin’s lymphoma
    h. Multiple myeloma
i. Myelodysplastic syndrome
j. Non-hodgkins lymphoma
k. Ovarian and testicular cancer
l. Sarcomas
m. Severe aplastic anemia
n. Severe combined immunodeficiency disease
o. Small-cell lung cancer
p. Other ______

13. What was the source of stem cells?
   a. Unrelated, matched donor
   b. Related, matched donor
   c. Cord blood donor

14. Please indicate any prior treatment regimens you may have received. Please check all that apply.
   a. Total body irradiation (TBI) full dose
   b. Total body irradiation (TBI) reduced/mini
   c. High dose chemotherapy
   d. Cranial irradiation
   e. Intrathecal chemotherapy
   f. Systemic chemotherapy (including hydroxurea, interferon, etc.)

15. Did you at any time develop graft versus host disease (GVHD)?
   a. No
   b. Yes; acute GVHD
   c. Yes; acute GVHD which developed into chronic GVHD
   d. Yes; late-onset chronic GVHD

16. If you developed chronic GVHD, please indicate which of the following systemic chronic GVHD medication you have taken or continue to take.
   a. Cyclosporine
   b. Tacrolimus
   c. Mycophenolate mofetil
   d. Glucocorticoids
   e. Other ______

Please answer yes or no to the following questions about your alcohol or prescription drug usage. If questions do not apply to you, please answer no.

17. Have you ever felt you should cut down on your drinking?
   a. Yes
   b. No

18. Have people annoyed you by criticizing your drinking?
   a. Yes
b. No

19. Have you ever felt bad or guilty about your drinking?
   a. Yes
   b. No

20. Have you ever had a drink first thing in the morning (as an “eye opener”) to steady your nerves or get rid of a hangover?
   a. Yes
   b. No

21. Have you ever taken over the prescribed amount of a prescription medication?
   a. Yes
   b. No
Appendix F: Demographics Questionnaire for HCT Partners

1. How old are you?

2. Are you
   a. Male
   b. Female

3. What is your ethnicity?
   a. Not Hispanic or Latino
   b. Hispanic or Latino

4. Some people identify themselves as belonging to one or more racial groups. Please indicate which of the following groups you belong to. Please check all that apply.
   a. White or Caucasian
   b. Black or African-American
   c. Hispanic or Latino
   d. American Native
   e. Alaskan Native
   f. Asian
   g. Pacific Islander
   h. Middle Eastern
   i. Other ______

5. How many years of education have you completed? (Completing high school or its equivalent = 12 years)

6. What is the highest level of education you have obtained?
   a. None
   b. High School Diploma
   c. GED
   d. Associates
   e. Bachelors
   f. Masters
   g. Doctorate
   h. Other (please specify)

7. What is your current marital status?
   a. Married
   b. Single
   c. Divorced
   d. Remarried
   e. Widowed
   f. Separated
   g. Living with partner (same sex)
h. Living with partner (opposite sex)

8. What is your current employment status? (Please check one)
   a. Full Time (>35 hrs/wk)
   b. Part Time (Regular hours)
   c. Part Time (Irregular hours)
   d. Unemployed, full time student
   e. Unemployed, part time student
   f. Unemployed due to health issues related to transplant
   g. Retired for transplant related issues
   h. Retired
   i. On disability for transplant related issues
   j. On disability for other reasons
   k. Military service

9. What is the current economic status of your household? (Please check one)
   a. We have barely enough to get by
   b. We have enough to get by, but no more
   c. We are solidly middle class
   d. We have plenty of “extras”
   e. We have plenty of “luxuries”
   f. Don’t know/unsure/prefer not to say

10. What is your annual household income? (Please check one)
    a. >$150,000
    b. $100,000-$149,000
    c. $75,000-$99,000
    d. $50,000-$74,000
    e. $25,000-$49,000
    f. $10,000-$24,000
    g. <$9,000
    h. Don’t know, or prefer not to say

Please answer yes or no to the following questions about your alcohol or prescription drug usage. If questions do not apply to you, please answer no.

11. Have you ever felt you should cut down on your drinking?
    a. Yes
    b. No

12. Have people annoyed you by criticizing your drinking?
    a. Yes
    b. No

13. Have you ever felt bad or guilty about your drinking?
    a. Yes
14. Have you ever had a drink first thing in the morning (as an “eye opener”) to steady your nerves or get rid of a hangover?
   a. Yes
   b. No

15. Have you ever taken over the prescribed amount of a prescription medication?
   a. Yes
   b. No
This questionnaire is an important part of providing you with the best health care possible. Your answers will help in understanding problems that you may have. Please answer every question to the best of your ability unless you are requested to skip over a question.

Please fill out these questions yourself, if at all possible. It’s important the information is from YOU.

I filled it out myself [ ] Someone else filled it out for me [ ]

<table>
<thead>
<tr>
<th>Over the last 2 weeks, how often have you been bothered by any of the following problems?</th>
<th>Not at all</th>
<th>Several Days</th>
<th>More than half of the days</th>
<th>Nearly every day</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Little interest or pleasure in doing things..............................................</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>b. Feeling down, depressed, or hopeless.........................................................</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>c. Feeling nervous, anxious or on edge.........................................................</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>d. Not being able to stop or control worrying.............................................</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How often in the past month did you:</th>
<th>Not at all</th>
<th>1-3 days</th>
<th>4-7 days</th>
<th>8-14 days</th>
<th>15-21 days</th>
<th>22-31 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Have trouble falling asleep........</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Wake up several times per night?</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Have trouble staying asleep (including waking far too early)?</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Wake up after your usual amount of sleep feeling tired and worn out?</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>
Appendix H: Test order and administration times

<table>
<thead>
<tr>
<th>Traditional Test</th>
<th>Time (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCFT</td>
<td>10</td>
</tr>
<tr>
<td>HVLT</td>
<td>5</td>
</tr>
<tr>
<td>Span</td>
<td>5</td>
</tr>
<tr>
<td>NART</td>
<td>5</td>
</tr>
<tr>
<td>SDMT</td>
<td>1.5</td>
</tr>
<tr>
<td>TMT</td>
<td>5</td>
</tr>
<tr>
<td>GPB</td>
<td>2</td>
</tr>
<tr>
<td>RCFT-R</td>
<td>2</td>
</tr>
<tr>
<td>HVLT-R</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>37.5</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CogState Task</th>
<th>Time (minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learn</td>
<td>5</td>
</tr>
<tr>
<td>One-Back</td>
<td>2</td>
</tr>
<tr>
<td>ISLT</td>
<td>5</td>
</tr>
<tr>
<td>Identification</td>
<td>2</td>
</tr>
<tr>
<td>GMLT</td>
<td>5</td>
</tr>
<tr>
<td>Detect</td>
<td>2</td>
</tr>
<tr>
<td>CPAL</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>26</strong></td>
</tr>
</tbody>
</table>
Appendix I: CogState Tutorial Script

The first part of testing will involve the use of this laptop. [Open welcome screen].

Testing will involve five separate tasks which will have separate directions associated with each.

Some parts of testing will make use of the mouse [Demonstrate mouse movement and icon clicking].

For other parts, you will be using the keypad on this laptop and specifically, these three keys [demonstrate “d,” “k” and “spacebar” keys].

Please take a moment to familiarize yourself with the laptop. [Provide 2 minutes for the participant to experiment with clicking the mouse and pressing keys].

This computerized portion of testing will take approximately 20 minutes. For each of the five tasks, there will be a separate set of directions that I will read you.

Please let me know when you are ready to begin.
Appendix J: Debriefing Sheet: Validation of the Adapted CogState Brief Battery in Hematopoietic Cell Transplant Patients

Thank you for participating in our study. The information you have provided is greatly appreciated and will be useful as we strive to learn more about the effects of hematopoietic cell transplant on thinking, memory and attention. It is our hope that information gained will be used to develop systems to help others in the future.

We understand that many aspects of your life have been affected by your transplant. If you are interested in any services or support to assist you as you continue to adjust, a list of local resources is outlined below.

U of M Blood & Marrow and Stem Cell Transplant Support Group
  - Meets every Monday, 1-2pm, in UH, 8A Family Lounge
  - For families and patients

  National Bone Marrow Transplant Link: nbmtlink.org

Bone and Marrow Transplant Information Network: bmtinfonet.org

The Leukemia and Lymphoma Society: leukemia-lymphoma.org
  - Online resources, support and information for families and patients

Again, we greatly appreciate your participation in our study.
Appendix K

Research study seeking volunteers

Are you a bone marrow or stem cell transplant survivor or spouse of someone who has undergone a stem cell transplant?

Take part in an important study examining the impact of bone marrow transplant on thinking, memory and attention!

The Behavioral Medicine Research Team at Eastern Michigan University is conducting a study on methods of evaluating how adult bone marrow or stem cell transplantation may affect memory, attention and thinking processes.

We encourage persons and their partners over the age of 18 years who have had a bone marrow or stem cell transplant within the past three years to sign up for participation!

Study participants will be asked to complete a series of brief questionnaires as well as memory and attention tasks on the computer and with a study administrator.

Participation will take approximately 30 minutes to complete.

The results of this study may benefit other transplant survivors through informing future investigations of cognitive functioning and rehabilitation programs.

To find out more about the study or sign up to participate, please contact the principle investigator, Bethany Gourley at bgourley@emich.edu or 734-487-4987.

This study has been approved by the Eastern Michigan University Human Subjects Review Committee. HSRC Approval # 110805
Appendix L: Personal Communication

CogState

David Darby <ddarby@cogstate.com>  
To: Bethany Gourley <bgourley@gmail.com>  
Cc: Paul Maruff <pmaruff@cogstate.com>  

Dear Bethany

It's great to hear from you again. I'm glad your research career is progressing.

If you are a student doing a study that is of interest to us, we would consider providing you with a coupon which waives the CogState Research fee (of $500 pa). You would need a new coupon if your research went beyond one year. To be considered for this, you would need to send us your protocol (preferably already IRB approved) and we'll go from there.

The results that are provided in CogState Research DataPoint accounts are:
1. Raw data - this is the decoded data file contents which has every keystroke and computer generated timepoint. It is explained by a help file, but would be a lot of work to do your thesis analyses if you only used this. It's provided so you can have full source data access if you need it to check our automated analyses etc.
2. Summary data - this is available as a spreadsheet containing all your accounts subjects' test results by task. It contains higher level data (ie means, std devs, accuracy, correct and error counts, etc). It has done a "smart" analysis, by which we mean it will exclude post-anticipatory correct responses for those tasks where this is important before computing the speed measures. This is something that we've found is appropriate since subjects who respond prematurely and then again once the stimulus appears tend to respond more slowly than when to the stimulus directly (probably related to having to disengage before re-engaging to the stimulus). In addition, the tasks themselves look for these, and reschedule extra trials if they are found (so a full complement of correct trial responses can be measured). In addition, the appropriate transformations of the speed and accuracy data are already performed to get these results (ie log 10 of each RT for which mean and std dev are computed, arcsine(sqrt()) of the hit rate etc). This data is usually the most appropriate for research studies, especially since most research studies includes controls to which the test condition subjects' performance is compared.
3. Clinical style reports - these include only results of a few of the tasks (Detection, Identification, One card learning, One back, GMLT errors) since these have sufficient age-based normative data and have been found useful in community screening or mass testing. These are used by some researchers to provide individual feedback (particularly in longitudinal studies where motivation and enthusiasm for testing is important for continued participation). These contains standardized scores (mean 100, sd 10) and standardized change scores for the subjects based on our normative data.

If you did wish to do your own comparison with control data, as long as we have some, we can share it with you.

I've cc'd Prof Paul Maruff on this email, as he is CSO, and is in charge of such facilitated research.

Kind regards

A/Prof David Darby  
Chief Medical Officer  
Em: ddarby@cogstate.com  
Web: www.cogstate.com

CogState Ltd. Lv 2, 255 Bourke Street Melbourne, Vic, 3000, Australia
On 14/06/2011, at 11:29 PM, Bethany Gourley wrote:

Dear Dr. Darby,

Last December, I met with you, Bruno Giordani and my research mentor, Flora Hoodin, with regards to using the CogState battery in bone marrow transplant patients. I hope that all has been well with you.

I have been working toward proposing my thesis project which will seek to evaluate the validity of CogState for use in bone marrow transplant patients. The study will be modeled after Maruff and colleagues' previous validation in 2009, so I plan to include the Brief Battery tests (Learn, Detect, Identify and One-Back). However, I would also like to include the CPAL, GMLT, and ISLT, as cognitive deficits in verbal learning and executive function are also often seen in bone marrow transplant patients.

I had a few logistical questions for you as I proceed with the study proposal.

Firstly, last winter, you suggested that I could potentially use CogState for a reduced cost. I was wondering if this was still the case and if so, how I might go about accessing the software.

Secondly, could you clarify what kind of data is provided within the Datapoint report (i.e. standardized score, raw scores, norms)? The reason I ask is I am considering comparing my clinical group mean scores to norms in order to calculate the non-overlap statistic, as previously performed by Maruff et al in establishing criterion validity. Because this is a thesis project, I am expected to run my own statistical analyses.

Therefore, I am hoping you would be willing to provide me with access to healthy adult norms for these tests.

Thank you very much in advance for your help. I look forward to hearing from you.

Best wishes,

Bethany Gourley

Clinical Psychological Doctoral Fellow

Eastern Michigan University

Appendix M: Personal communication
Appendix N: Normative data for Adapted CogState Brief Battery tasks
<table>
<thead>
<tr>
<th>TASK</th>
<th>Source</th>
<th>Age Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Age Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Age Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Age Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>DET</td>
<td>Normative DB</td>
<td>18-34</td>
<td>120</td>
<td>2.46</td>
<td>0.09</td>
<td>35-50</td>
<td>400</td>
<td>2.49</td>
<td>0.1</td>
<td>51-60</td>
<td>461</td>
<td>2.53</td>
<td>0.11</td>
<td>61-70</td>
<td>560</td>
<td>2.54</td>
<td>0.11</td>
</tr>
<tr>
<td>IDN</td>
<td>Normative DB</td>
<td>1202</td>
<td>2.66</td>
<td>0.08</td>
<td>395</td>
<td>2.7</td>
<td>0.07</td>
<td>368</td>
<td>2.71</td>
<td>0.08</td>
<td>724</td>
<td>2.72</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OCL</td>
<td>Normative DB</td>
<td>683</td>
<td>1.05</td>
<td>0.13</td>
<td>212</td>
<td>1.01</td>
<td>0.12</td>
<td>206</td>
<td>1.02</td>
<td>0.11</td>
<td>575</td>
<td>1.02</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ONB</td>
<td>Normative DB</td>
<td>636</td>
<td>1.37</td>
<td>0.14</td>
<td>101</td>
<td>1.34</td>
<td>0.19</td>
<td>101</td>
<td>1.38</td>
<td>0.15</td>
<td>570</td>
<td>1.37</td>
<td>0.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMLT</td>
<td>Normative DB</td>
<td>853</td>
<td>40.88</td>
<td>15.34</td>
<td>341</td>
<td>55.36</td>
<td>22.42</td>
<td>298</td>
<td>55.7</td>
<td>28.78</td>
<td>329</td>
<td>62.29</td>
<td>23.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISLT</td>
<td>Normative DB</td>
<td>111</td>
<td>25.17</td>
<td>4.3</td>
<td>243</td>
<td>22.79</td>
<td>5.41</td>
<td>257</td>
<td>26.54</td>
<td>4.35</td>
<td>124</td>
<td>25.98</td>
<td>4.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISRL</td>
<td>Normative DB</td>
<td>259</td>
<td>9.25</td>
<td>3.02</td>
<td>268</td>
<td>8.29</td>
<td>3.16</td>
<td>224</td>
<td>8.21</td>
<td>2.36</td>
<td>141</td>
<td>9.32</td>
<td>2.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TASK</th>
<th>Source</th>
<th>18-34</th>
<th>Age Group</th>
<th>39-49</th>
<th>Age Group</th>
<th>50-59</th>
<th>Age Group</th>
<th>60-69</th>
<th>Age Group</th>
<th>70-79</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPFAL</td>
<td>8 probes</td>
<td>Normative DB</td>
<td>71</td>
<td>23.00</td>
<td>37.37</td>
<td>35</td>
<td>54.94</td>
<td>40.24</td>
<td>23</td>
<td>62.28</td>
</tr>
</tbody>
</table>
Appendix O. Effect sizes and 95% confidence intervals for correlations between CogState and traditional tests

Fig. 3. Effect sizes and their 95% confidence intervals (95% CIs) for correlations between CogState tasks and traditional neuropsychological tests for $r$'s $\geq 0.30$