Victoria Symptom Validity Test performance in acute severe traumatic brain injury: Implications for test interpretation

Stephen N. Macciocchi*, Ronald T. Seel, Amy Alderson, Robert Godsall

Shepherd Center, 2020 Peachtree Road, Atlanta, GA, United States
Accepted 12 June 2006

Abstract

Effort testing has become commonplace in clinical practice. Recent research has shown that performance on effort tests is highly correlated with performance on neuropsychological measures. Clinical application of effort testing is highly dependent on research derived interpretive guidelines. The Victoria Symptom Validity Test (VSVT) is one of many measures currently used in clinical practice. The VSVT has recommended interpretive guidelines published in the test manual, but the samples used in developing interpretive guidelines are small and heterogeneous and concern has been expressed regarding high false negative rates. In this study, a homogeneous sample of acute, severely brain injured persons were used to assess the sensitivity of the VSVT. Results confirmed that acute, severely brain injured persons (N = 71) perform very well on the VSVT. The severe brain injury population is 99% likely to have between 44.1 and 46.8 correct VSVT Combined Score responses. While the VSVT was insensitive to memory dysfunction, the presence of severe visual perceptual (Benton Visual Form Discrimination Score < 21) and verbal fluency (Controlled Oral Word Association Score < 15) deficits predicted poor performance on the VSVT. These results provide further evidence that performance expectations currently incorporated in the VSVT manual interpretative criteria are too conservative. Empirically based alternative criteria for interpreting VSVT Combined Scores in the TBI population are presented.

Keywords: Symptom validity; Victoria Symptom Validity Test; Traumatic brain injury

1. Introduction

For some time, neuropsychologists have understood that neurocognitive test performance can be influenced by factors other than central nervous system trauma and disease, including age, education, medical disorders, persistent pain, emotional distress, and medication (Lishman, 1997; Perna, Williams, Durgin, Geller, & Allen, 2000). More recently, clinicians and researchers have become particularly concerned about the impact of effort on neuropsychological test performance (Bianchini, Mathias, & Greve, 2001; Lynch, 2004; Rees, Tombough, & Boulay, 2001; Slick, Tan, Strauss, & Hultsch, 2004). Concerns about suboptimal effort have been fueled by the exponential increase in medicolegal consultations, but threats to test validity related to effort have always been a concern for most neuropsychologists...
In general, most effort tests focus on memory functioning and require examinees to learn and recall various types of information over variable time periods. In current test paradigms, stimuli to be learned and retained vary and may involve pictures of objects as in the Test of Memory Malingering [TOMM] (Tombaugh, 1996), digit sets as in the Portland Digit Recognition Test [PDRT] (Binder, 1993; Binder & Willis, 1991), or word lists as used in the Word Memory Test [WMT] (Green, 2000). In each case, the test has been designed to allow for highly successful performance, even when persons have neurologic disorders that significantly compromise memory functioning. When performance on effort tests is impaired, neuropsychological test performance also is very likely to be disproportionately impaired, presumably independent of neurologic trauma-disease (Green, Rohling, Lees-Haley, & Allen, 2001).

Assessing the impact of effort on test performance can be multi-faceted endeavor typically involving a number of data points and interpretive strategies (Lynch, 2004; Slick et al., 2004). Clinical observation, effort testing and utilization of other techniques such as Bayesian Model Averaging (BMA) may ultimately be necessary to fully evaluate the validity of test data obtained during neuropsychological examinations, particularly in cases where compensation is an issue (Millis & Volinsky, 2001). While multiple measures and strategies may be necessary to adequately assess effort, clinicians must have an adequate empirical basis for each interpretive strategy.

The Victoria Symptom Validity Test (VSVT) (Slick, Hopp, Strauss, & Spellacy, 1996) is an effort measure that focuses on memory functioning and requires examinees to learn and recall digits. The test manual (Slick, Hopp, Strauss, & Thompson, 1997) indicates that valid performances fall between 30 and 48 correct responses out of a possible 48 test items. Questionable performance falls between 18 and 29 correct responses and “invalid” performance is 17 or fewer correct responses. Valid mean latencies described in the test manual range from 1.67 s on “Easy” items to 2.68 s on “Hard” items. Invalid mean latency times ranged from 3.40 s on “Easy” items to 4.70 s on “Hard” items.

Interpretive guidelines for VSVT scores were largely based on binomial probability theory. While theoretically a valid approach, concern has been expressed that these statistic based odds for determining invalid performance are too conservative in clinical practice and yield too many false negatives. Using the interpretive guidelines as well as data reported in the VSVT manual, approximately 60% of persons who were asked to feign cognitive impairments on the VSVT had false negative “valid” scores while only 5% who feigned cognitive impairment had true positive “invalid” scores. A study of persons with epilepsy found that persons seeking compensation performed significantly worse on the VSVT than non-compensation seeking individuals (Grote et al., 2000). Nonetheless, mean performances of persons seeking compensation (mean = 39.7) fell in the valid range according to interpretive guidelines.

In a recent review, Bianchini et al. (2001) recommended establishing population specific databases for all effort tests. At the current time, the VSVT has not been extensively studied in severely brain injured samples. Recent research has shown the VSVT to be insensitive to significant memory impairment (Slick, Tan, Strauss, Mateer, Harnadek, & Sherman, 2003), but there is a need to examine VSVT performance in a larger sample of persons with acute cerebral dysfunction secondary to acquired brain injury to identify whether other cognitive deficits might impact VSVT performance. Data from persons with acute, severe brain injury would provide a basis for supporting current interpretive guidelines or establishing new guidelines.

Accordingly, we administered the VSVT and a neuropsychological test battery to persons who recently emerged from delirium following brain injury. The study objectives were to:

1. describe the pattern of VSVT performance for persons with acute severe brain injuries;
2. examine the impact of demographic, brain injury severity, and psychological factors on VSVT test scores;
3. identify specific neurocognitive impairments associated with VSVT performance;
4. confirm existing guidelines, or suggest new guidelines and identify “Questionable” and “Likely Invalid” effort utilizing a severe, acute brain injury sample.

Our primary hypothesis was that persons with acute severe brain injury would perform above the currently recommended cut-off (30) for a “valid” profile. We also hypothesized that persons with acute TBI would demonstrate longer response latencies than persons profiled in the VSVT manual. Lastly, based on the similarity of test demands,
we hypothesized that patient performance on the Digit Span Test would be most highly correlated and predictive of VSVT Total Score performance.

2. Method

2.1. Setting

The study was conducted at a specialized, CARF and JCAHO accredited neurorehabilitation hospital located in Atlanta, Georgia. The hospital has a comprehensive Acquired Brain Injury Program which includes an acute inpatient rehabilitation unit, an outpatient day program and a residential program. Interdisciplinary rehabilitation teams consisting of healthcare professionals from physiatry, nursing, and neuropsychology as well as physical, occupational, speech and recreational therapy provide care for brain injured persons on a daily basis. Standardized injury assessment, treatment, and outcome data are routinely collected and brain injured persons are assigned to treatment modules based on carepaths using a quantitative severity index that identifies specific treatment goals and expected outcomes. All persons in the current study were admitted to the acute neurorehabilitation inpatient program secondary to a documented history of brain injury and in some cases co-morbid extracranial injury. Neuropsychological testing was conducted as part of a standardized clinical assessment completed prior to discharge from the inpatient program.

2.2. Participants

Persons participating in the study \((n=71)\) were consecutive referrals assessed over a 9-month period who evidenced cognitive and functional deficits secondary to brain injury sufficient to merit an extended interdisciplinary rehabilitation admission based on contemporary medical and insurance criteria. All persons admitted with a primary diagnosis of a traumatic brain injury were considered for inclusion. Persons with a primary diagnosis of cerebral vascular accident \((n=9)\) or anoxic/other non-traumatic brain injury \((n=5)\) were excluded. Persons with severe TBI who did not emerge from their confusion (e.g. Rancho 4 or less), could not follow commands, or who had profound visual or motor dysfunction that precluded administration of neuropsychological tests \((n=21)\) were also excluded from participation. No participants were actively involved in litigation at the time of enrollment and testing.

Demographic analyses revealed that participants were predominantly male (76%). Mean age at time of evaluation was 28.1 (S.D. = 11.9) with a median of 24 and a range of 15–55 years. Participants mean education was 12.3 years (S.D. = 2.7) with a median of 12 and a range of 5–20 years. For those with data available \((n=44)\), the mean Glasgow Coma Scale (GCS) score at systems admission was 7.6 (S.D. = 3.9) with a median of 7. Neuroimaging documented that 34% of participants evidenced diffuse axonal injury, while 20% had a subarachnoid hemorrhage and 14% had an intraventricular hemorrhage. Neuropsychological and symptom validity testing was conducted a mean of 43.4 days post-injury (S.D. = 23.9) with a median of 38 and a range of 6–116 days. On the day of testing, patients’ mean number of errors on the Galveston Orientation and Amnesia Test (GOAT) was 18.3 (S.D. = 20.2) with a median of 12. When tested, a significant number of participants (20%) had 25 or greater error points on the GOAT. Mean score on the Token Test at the time of testing was 40.6 (S.D. = 4.5) with a median score of 42.

2.3. Data collection

Informed consent was obtained from either the participant or their legally authorized representative. Research staff collected injury and demographic data from medical records. Primary injury information included neuroimaging findings and GCS at time of initial medical systems admission. Participants were scheduled for neuropsychological testing within 1 week of becoming oriented in all spheres. At that time, a clinical neuropsychologist interviewed the patient and a psychometrist blinded to study hypotheses and methodology administered the neuropsychological and effort tests in a controlled environment. Tests and inventories administered to participants included the Beck Depression Inventory (Beck, Steer, & Brown, 1996), the Benton Visual Form Discrimination Test—BVFD (Benton, Sivan, Hamsher, Varney, & Spreen, 1994), the Controlled Oral Word Association Test—COWA (Benton & Hamsher, 1989), the Galveston Orientation and Amnesia Test (Levin, O’Donnell, & Grossman, 1979), the Rey Auditory Verbal Learning Test (Rey, 1964), Symbol Digit Modalities Test—SDMT (Smith, 1982), the Token Test (Boller & Vignolo,
Due to the clinical nature of the evaluation, some neuropsychological tests were not administered or were discontinued due to patients’ cognitive impairments or fatigue, resulting in incomplete quantitative data. Thus, test completion code variables were generated for some tests. Scores were calculated for each test and the neuropsychology and research staff performed quality control checks on the data. Following assessment and scoring, a research coordinator organized and entered the clinical and test data into a data file. All data entry and analyses were conducted using the Statistical Package for the Social Sciences (SPSS) Version 11.5. Identifying information was removed from the research data file to protect patient privacy in compliance with the Health Insurance Portability and Accountability Act of 1996 (HIPAA) regulations.

2.4. Data analyses

Descriptive statistics including means, medians, and standard deviations were compiled for quantitative demographic, injury, psychological, and neuropsychological measures. Given the importance of understanding the population performance of brain injured persons on the VSVT, standard error of the mean (S.E.) and confidence intervals (CI) were computed on all VSVT scores and response latencies. Confidence intervals were set at the 99th percentile and computed based on a t-distribution sample correction parameter of 2.66 that conservatively used 60 degrees of freedom (Portney & Watkins, 2000, pp. 394–395).

To examine the association of demographic, injury, and psychological factors with VSVT scores and latencies, multiple approaches were required to meet the statistical assumptions of comparing variables that used different rating schemes (categorical versus quantitative) or had normal versus non-normal distributions. The impact of categorical variables (e.g. gender and presence of DAI, SAH, and IVH) on normally distributed VSVT scores and latencies was examined using one-way analysis of variance (ANOVA) (Daniel, 1995). The impact of categorical variables on non-normally distributed VSVT scores and latencies was examined using Kolmogorov–Smirnov tests (K-S test) (Porter & Hamm, 1986). The K-S test was utilized as a non-parametric alternative to the t-test in order to address rank-ordered, ordinal data or non-normally distributed data from two independent samples. The K-S test generates a positive or negative z score, with higher absolute values demonstrating greater differences between two samples. Pearson r correlations were used to document the association among quantitative variables (e.g. mean age at time of evaluation; GCS score at initial systems admission; mean number of days from injury to evaluation; GOAT Total Errors; Token Test Total Score; and BDI Total Score) with normally distributed VSVT scores and latencies. Where bivariate comparisons involved at least one non-normally distributed quantitative variable, Spearman rho ($\rho$) rank-ordered correlations were calculated. Alpha was set at .01 for these analyses in order to minimize the probability of making Type I errors.

With regard to selecting variables for the two multi-variate regression analyses, bivariate analyses were first conducted to examine the relationship between neuropsychological and VSVT performance (items correct and latencies). Where bivariate analyses revealed highly correlated predictor variables ($\rho$ or $r \geq .60$), one of each pair was deleted. To test for linearity and homoscedasticity, bivariate scatter plots were examined. A stepwise regression analysis was used to identify those variables that added a significant amount of variance to the model. Final models included only statistically significant predictors ($p < .05$). For the two multi-variate regression models, $R^2$ indicated the proportion of variance contributed by predictor variables to each outcome measure.

To identify logical cut-off scores for VSVT Combined Scores, the following procedure was utilized. Predictors of VSVT Combined Scores from the regression analysis were selected as criterion variables. Stem-and-leaf scatter plots were then calculated in which patients’ VSVT Combined Scale scores were mapped to their criterion variable performance. Cut-off scores were derived based on a decision rule that emphasized the minimization of false positives for Questionable effort.
Table 1

<table>
<thead>
<tr>
<th>VSVT scores and response latencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>Easy items</td>
</tr>
<tr>
<td>#Correct</td>
</tr>
<tr>
<td>Response latency (s)</td>
</tr>
<tr>
<td>Hard items</td>
</tr>
<tr>
<td>#Correct</td>
</tr>
<tr>
<td>Response latency (s)</td>
</tr>
<tr>
<td>Combined Scores</td>
</tr>
<tr>
<td>Total #correct</td>
</tr>
<tr>
<td>Easy minus Hard</td>
</tr>
</tbody>
</table>

Note. S.D.: standard deviation; S.E.: standard error of the mean; CI: confidence interval.

3. Results

3.1. Pattern of VSVT scores

Descriptive statistics were calculated to analyze brain injured persons’ VSVT performance (see Table 1). Persons with acute TBI performed well on the VSVT Easy items ($M = 23.38; S.D. = 2.01$) with almost 96% scoring 22 correct or higher. Likewise, participants performed well on the VSVT Hard items ($M = 22.11; S.D. = 2.80$). Seventy-two percent scored 22 correct or higher with somewhat greater performance variability than on the Easy items. A frequency distribution analysis of VSVT Combined Scores revealed a negatively skewed sample with considerable kurtosis (see Fig. 1). Nearly two-third of the sample score had a Combined Score of 47 or 48 correct. Only 3 cases had less than 38 correct responses.

On average, the response latency for the VSVT Easy items was 3.38 s (S.D. = 1.83). Eighty-seven percent of participants had Easy item mean response latencies between 1.06 and 4.66 s. Participants required more time to complete VSVT Hard items, averaging 4.62 s per item (S.D. = 2.25). Ninety percent of the sample had mean Hard item response latencies between 1.05 and 6.89 s.

Bivariate correlations were then calculated to examine the relationship between VSVT test scores and response latencies. The VSVT Hard item scores ($r_{rho} = .984, p < .001$) were more highly associated with the VSVT Combined Scores than the VSVT Easy item scores ($r_{rho} = .615, p < .001$). VSVT Easy and Hard item scores were moderately

![Fig. 1. Frequency distribution of VSVT Combined Scores ($n = 71$).](image-url)
correlated with each other \( (S_{\rho} = .519, p < .001) \). VSVT Easy and Hard item response latencies were very highly and positively correlated with each other \( (r = .890, p < .001) \), suggesting a close to perfect linear relationship. However, neither VSVT Easy item scores and response latencies \( (S_{\rho} = -.288, p = .015) \) nor VSVT Hard item scores and response latencies \( (S_{\rho} = -.239, p = .046) \) were significantly correlated. Likewise, VSVT Easy item \( (S_{\rho} = -.248, p = .038) \) and Hard item response latencies \( (S_{\rho} = -.256, p = .032) \) had little relationship to VSVT Combined Scores.

### 3.2. Demographic, brain injury severity, psychological factors, and VSVT scores

K-S tests revealed that gender \( (p = .386) \) and neuroimaging findings including DAI \( (p = 1.000) \), SAH \( (p = .834) \), and IVH \( (p = .281) \) were not significantly related to VSVT Combined Scores. Age at time of evaluation \( (p = .348) \) and GOAT Total Errors \( (p = .120) \), were not significantly correlated to VSVT Combined Scores. Two variables were found to be modestly correlated with VSVT Combined Scores—GCS score at initial medical systems admission \( (S_{\rho} = .407, p < .01) \) and Token Test Total Score \( (S_{\rho} = .307, p = .010) \), while years of education \( (S_{\rho} = .300, p = .011) \) and number of days from injury to neuropsychological testing \( (S_{\rho} = -.296, p = .012) \) approached statistical significance. With regard to psychological functioning, BDI Total scores were inversely associated with VSVT Combined Scores \( (S_{\rho} = -.381, p < .001) \).

ANOVA revealed that gender \( (p = .792) \) and neuroimaging findings including DAI \( (p = .586) \), SAH \( (p = .409) \), and IVH \( (p = .985) \) were not significantly related to VSVT Hard item response latencies. Age at time of evaluation \( (p = .596) \) and GOAT Total Errors \( (p = .029) \) were also not significantly correlated to VSVT Hard item latencies. Years of education \( (r = -.412, p < .01) \), GCS score at initial medical systems admission \( (r = -.394, p < .01) \), and Token Test Total Score \( (r = -.374, p < .01) \) were significantly correlated with VSVT Hard item latencies. Number of days from injury to neuropsychological testing \( (r = .303, p = .011) \) approached statistical significance. With regard to psychological functioning, BDI Total scores were not correlated with Hard item latencies \( (p = .207) \).

### 3.3. Predictors of VSVT scores

To identify specific acute cognitive impairments that impacted VSVT performances, stepwise multiple regression analyses were conducted to identify neuropsychological and patient factors predictive of VSVT Combined Scores. Bivariate correlations were conducted to identify neuropsychological test performances that were most associated with VSVT Combined Scores. Twelve neuropsychological test scores were significantly correlated \( (p < .01) \) with VSVT Combined Scores: COWA \( (S_{\rho} = .567) \); Blocks Correct \( (S_{\rho} = .536) \); SDMT Written \( (S_{\rho} = .529) \); SDMT Oral \( (S_{\rho} = .528) \); WMS-R Visual Reproduction I \( (S_{\rho} = .498) \); BVFD \( (S_{\rho} = .473) \); Digit Span Backwards Score \( (S_{\rho} = .467) \); Trails A \( (S_{\rho} = -.415) \); Trails B \( (S_{\rho} = -.403) \); WMS-R Visual Reproduction II \( (S_{\rho} = .356) \); RVLT Delayed Recall \( (S_{\rho} = .318) \); and RVLT Trials 1–5 Total Score \( (S_{\rho} = .309) \). Several neuropsychological predictor variables were highly correlated \( (r \geq .600) \) and one of each pair was removed from further analysis. Six independent variables were identified as potential predictors of VSVT Combined Scores and entered into the regression equation including COWA, SDMT Oral, BVFD Total Score, Digit Span Backwards Score, Trails A, and BDI Total Score. Table 2 displays the \( R^2 \), \( R^2 \) change, \( F \) change, and statistical significance \( (p) \) of each predictor variable. Two of the six entered independent variables (BVFD and COWA) contributed significantly to the prediction model for VSVT Combined Scores. The two variables in combination contributed 36.3% in shared variance.

A second stepwise multiple regression analysis was conducted to identify predictors of VSVT Hard item response latencies. Twelve neuropsychological test scores were also significantly correlated with VSVT Hard item latencies at \( p < .01 \) including SDMT Oral \( (r = -.625) \); Trails A \( (r = .601) \); Trails B \( (r = .586) \); SDMT Written \( (r = -.573) \); Blocks

<table>
<thead>
<tr>
<th>Model</th>
<th>( R^2 )</th>
<th>( R^2 ) change</th>
<th>( F ) change</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>BVFD Total</td>
<td>.275</td>
<td>-</td>
<td>24.33</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>BVFD + COWA</td>
<td>.363</td>
<td>.087</td>
<td>8.65</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Excluded variables: Digits Backward \( (p = .884) \), Symbol Digit Oral \( (p = .741) \), Trails A \( (p = .213) \), and BDI Total Score \( (p = .144) \).
Table 3
Regression model for VSVT Hard item response latencies

<table>
<thead>
<tr>
<th>Model</th>
<th>$R^2$</th>
<th>$R^2$ change</th>
<th>$F$ change</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable: VSVT Hard item response latency (s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trails A Seconds</td>
<td>.361</td>
<td>–</td>
<td>36.72</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Trails A + SDMT Oral</td>
<td>.455</td>
<td>.094</td>
<td>11.03</td>
<td>.001</td>
</tr>
</tbody>
</table>

Excluded variables: COWA ($p = .896$), BVFD ($p = .477$), years of education ($p = .194$), and WMS-R Logical Memory I ($p = .120$).

Correct ($r = -.456$); WMS-R Logical Memory I ($r = -.437$); WMS-R Visual Reproduction I ($r = -.406$); RVLT Trials 1–5 Total Score ($r = -.389$); BVFD ($r = -.385$); COWA ($r = -.385$); RVLT Delayed Recall ($r = -.366$); and WMS-R Logical Memory II ($r = -.347$). Several neuropsychological predictor variables were highly correlated ($r \geq .600$) and one of each pair was removed from further analysis. Six independent variables were identified as potential predictors of VSVT Hard item latencies and entered into the regression equation: SDMT Oral, Trails A, WMS-R Logical Memory I Score, years of education, BVFD, and COWA. Table 3 displays the $R^2$, $R^2$ change, $F$ change, and statistical significance ($p$) of each predictor variable. Two of the six entered independent variables (Trails A Seconds and SDMT Oral Score) contributed significantly to the prediction model for VSVT Hard item latencies. The two variables in combination contributed 44.5% in shared variance.

Based on the regression results, the BVFD and COWA scores were further evaluated to determine what level of visual perceptual and verbal initiation impairments lead to poor performance on the VSVT and would impact invalid cut-off scores. Stem-and-leaf plots were calculated to map BVFD scores and COWA scores to VSVT Combined Scores (see Fig. 2). Patients ($n = 52$) who scored $\geq 21$ on the BVFD and $\geq 15$ on the COWA test had a median VSVT score (solid black line) of 48 with the 25th percentile and above (red-shaded area) scoring 47 or greater and with a normal range ($T$-brackets $= 2S.D.$) of 46 or greater. Four outliers scored between a 41 and 45 and no person scored under 41. Patients ($n = 13$) who either had a BVFD score $< 21$ or COWA test score $< 15$ had a median VSVT score of 44 with the 25th percentile and above scoring 42 or greater and all patients scoring 38 or greater. When significantly impaired performances were observed on both the BVFD and COWA tests ($n = 5$), performance dropped significantly and varied widely with a median score of 30 and a range of 27–47.

Fig. 2. Stem-and-leaf plots of VSVT Combined Score in relation to acute severe TBI patient performance on BVFD and COWA.
4. Discussion

4.1. Pattern of VSVT performance for persons with acute severe brain injuries

Our primary hypothesis was that persons with acute severe brain injury would perform well above the VSVT manual recommended cut-off (30) for a “valid” profile. The results from our study confirm this hypothesis, indicating that the acute severely brain injured population is 99% likely to have a group mean between 44.12 and 46.86 correct on the VSVT Combined Score. This finding suggests that the acute severely brain injured population as well as those with lesser brain injuries can be expected to perform well on the VSVT with a very high degree of confidence. The extent of participants’ ability to perform well on the VSVT even when in the acute phase of a severe brain injury provides further evidence that the lower threshold of “valid” performance recommended in the VSVT manual is too conservative.

We also hypothesized that persons with acute TBI would demonstrate longer response latencies than persons profiled in the VSVT manual. The results from our study indicated that latencies for responding were uniformly increased and closely approximated performance designated as “invalid” by the test manual. In general, severely brain injured persons take longer to accurately respond than is indicated by classification guidelines in the test manual. In addition, correct responses and latencies were not significantly related. As such, when an examinee has a history of severe brain injury and performs well on the VSVT, slow response latencies should not be used as an indicator of suboptimal effort.

4.2. Impact of demographic, brain injury severity, and psychological factors

Severely brain injured person’s ability to correctly respond to VSVT items was not significantly associated with age, gender, education, neuroimaging findings, and GOAT score. Admission GCS score was modestly, but significantly correlated with VSVT performance. Depression was also modestly associated with poorer performance on the VSVT. However, neither depression nor admission GCS score were predictive of VSVT Combined Scores, suggesting that other factors play predominate mediating roles in differential performance on the VSVT. Thus, the VSVT appears to be insensitive to demographic and injury factors, which increases its generalizability to the broader traumatic brain injury population.

4.3. Specific neurocognitive impairments associated with VSVT performance

Thirteen neuropsychological test scores were significantly correlated with VSVT performance, but consistent with the test design the VSVT appears to be less sensitive to memory dysfunction than other cognitive functions. Regression analyses showed that performance on the BVFD and COWA were the best predictors of VSVT performance, which disconfirmed our hypothesis about Digit Span performance being the best predictor VSVT scores. One should note that BVFD scores <21 and COWA test scores <15 are generally well below the 1st percentile based on general population norms and represent very severely impaired performance. Likewise, vision played a considerable role in slow response latencies on the VSVT with primary predictors being visual motor dysfunction and cognitive processing speed with a visual component. Thus, persons with documented severe TBI and concomitant severe visual and initiation impairments may be at greater risk for having legitimate difficulty responding accurately to questions.

4.4. Suggested guidelines for “Questionable” and “Likely Invalid” effort

Assessing effort and the impact of effort on neuropsychological test performance is a multi-factorial process. Use of formal effort tests provides one index of neuropsychological test validity. Data from the current study clearly indicate that severely brain injured persons can perform well on the VSVT. A primary concern of the VSVT authors in establishing cut-off scores was to prevent false positive diagnoses because “an inaccurate diagnosis of malingering can be devastating to the individual”. The final question of the current study then becomes can less conservative interpretive criteria be empirically recommended for the VSVT that prevent false positive diagnoses, and improve true positive and false negative rates associated with the current interpretive guidelines?

We compiled VSVT Combined Score data for controls (n=95) and feigning individuals (n=43) from Table 3 of the VSVT manual along with VSVT Combined Score data from this study’s acute severe TBI group with intact visual perceptual and verbal initiation abilities (n=52). We then empirically set VSVT Combined Score interpretation
Table 4
Suggested VSVT Combined Score interpretive criteria for persons with TBI aged 18–55

<table>
<thead>
<tr>
<th>Interpretation criteria</th>
<th>Normal controls (n = 95) (%)</th>
<th>Feigning group (n = 43) (%)</th>
<th>Acute severe TBI* (n = 52) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likely Valid (44–48)</td>
<td>100b</td>
<td>10c</td>
<td>96b</td>
</tr>
<tr>
<td>Questionable (38–43)</td>
<td>0</td>
<td>15</td>
<td>4</td>
</tr>
<tr>
<td>Likely Invalid (37 or less)</td>
<td>0d</td>
<td>75e</td>
<td>0d</td>
</tr>
</tbody>
</table>

a Intact vision and verbal initiation abilities.
b True negative.
c False negative.
d False positive.
e True positive.

criteria as Likely Valid = 44–48, Questionable = 38–43, and Likely Invalid = 37 or less (see Table 4). The proposed interpretation criteria yielded no false positives diagnoses. Further, the number of true positive diagnoses (Likely Invalid for the feigning group) improved from 5% using the manual guidelines to 75% using the proposed guidelines. Lastly, the number of false negative diagnoses (Likely Valid in feigning group) reduced from 60% using the manual guidelines to 10%.

4.5. Limitations and future research

Several caveats should be considered. We focused on traumatically injured persons due to the high base rate of TBI-related personal injury claims. As such, our data may not be directly applicable to anoxic, metabolic or stroke disorders. Secondly, our sample was young and the effect of age is generalizable to the range of 16–55. Finally, we excluded patients who could not follow commands and were unable to take any tests, but persons with catastrophic impairment are rarely sent for extended neuropsychological assessment or forensic examinations. Future research is required to replicate findings from this study as well as examine VSVT performance in anoxic, metabolic, stroke, and other neurological populations.

5. Conclusions

The VSVT is insensitive to memory problems in severe brain injury and as such can provide an index of performance believed to reflect the complex construct of effort. Persons with known or suspected histories of brain injury who have even minimally functional visual perceptual abilities and verbal fluency functioning should perform extremely well on the VSVT. Finally, our data indicate that use of existing VSVT interpretive strategies merits modification if false negative decisions regarding neuropsychological test validity are to be avoided.

Acknowledgements

The authors wish to thank the Shepherd Center case managers and neuropsychological support staff for their cooperation and collaboration with the collection of data for this project. Also, the authors wish to thank James Mullin, Krista Wild, and Nicole Thompson for their assistance in preparing the data for analysis.

References


