The Medical Symptom Validity Test in the evaluation of Operation Iraqi Freedom/Operation Enduring Freedom soldiers: A preliminary study

Kriscinda A. Whitney a,b,*, Polly H. Shepard b,c, Amanda L. Williams d, Jeremy J. Davis e, Kenneth M. Adams f

a Richard L. Roudebush Veterans Affairs Medical Center, Indianapolis, IN, USA
b Indiana University School of Medicine, Indianapolis, IN, USA
c Professional Psychological Services, Indianapolis, IN, USA
d Ball State University, Muncie, IN, USA
e University of Indianapolis, Indianapolis, IN, USA
f VA Ann Arbor Healthcare System and University of Michigan Health System, Ann Arbor, MI, USA

Accepted 13 March 2009

Abstract

The clinical utility of the Medical Symptom Validity Test (MSVT) for soldiers returning from service in Operation Iraqi Freedom or Operation Enduring Freedom was preliminarily investigated through retrospective chart review. Results showed that 17%, or 4 of 23, Operation Iraqi Freedom/Operation Enduring Freedom patients at a Polytrauma Network Site (Level 2), performed below cut-offs on the MSVT. On “easy” subtests of the MSVT, the group of individuals who failed the MSVT performed significantly worse than the group of individuals who passed. However, there were no significant group differences on the “hard” subtests of the MSVT. When the profiles of individuals who failed the MSVT were examined, none of them met the criteria for the Dementia Profile. These preliminary findings and additional test data supported the conclusion that participants who failed the MSVT were exhibiting diminished symptom validity, suggesting that the specificity of the MSVT was 100%.

Keywords: Neuropsychology; Malingering; Military Veterans; Traumatic brain injury; Test validity; Memory

Introduction

Many Operation Iraqi Freedom/Operation Enduring Freedom soldiers have been exposed to blasts from improvised explosive devices or rocket-propelled grenades. Although specific data regarding the prevalence of blast exposure and possible mild traumatic brain injury (MTBI) among Operation Iraqi Freedom/Operation Enduring Freedom soldiers are surprisingly scarce, two recently published investigations do provide preliminary information on the topic. Specifically, results of a recent survey of 2,525 U.S. Army infantry soldiers who completed a 1-year deployment to Iraq showed that 124 (4.9%) reported injuries with loss of consciousness (LOC) and 260 (10.3%) reported injuries with altered mental status, but no LOC (Hoge, McGurk, Thomas, Cox, Engel, & Castro, 2008). Similarly, in a separate cross-sectional survey of military personnel following deployments to Iraq and/or Afghanistan, it was found that about 12% (n = 275) of 2,235 respondents reported a history consistent with MTBI (Schneiderman, Braver, & Kang, 2008). Of those 275 persons, 103 (4.6% of the total sample) reported having experienced a blast injury, 32 (1.4% of the total sample) of whom reported blast as their sole mechanism of injury. The data suggested that veterans exposed to only a blast and not to other injury mechanisms, such as a fall or motor vehicle crash, were 6.5 times more likely to experience an MTBI than veterans not exposed to any injury mechanism.
Whereas the aforementioned data suggest that blast exposure alone increases the risk for MTBI, they tell us little about the nature of blast injury effects. It has been proposed that the effects of blasts fall into four categories: primary (direct effects of pressure), secondary (effects of projectiles), tertiary (effects of structural collapse and of persons being thrown by the blast wind), and quaternary (burns, asphyxia, and exposure to toxic inhalants; DePalma, Burris, Champion, & Hodgson, 2005). However, although the primary effects of blasts are known to cause certain injuries, such as pulmonary damage and rupture of tympanic membranes, the latter of which may mimic persisting symptoms of MTBI, there are not yet published research data about the association between MTBI and primary blast effects in humans. Although research data is lacking, assessment and treatment of injured soldiers continues in the clinical setting, where, regardless of blast injury category, evaluation of symptoms of MTBI is a necessity.

With regard to the specifics of assessment in persons with MTBI, the current literature suggests that symptom validity testing is a vital evaluation component in this population. In fact, results of a recent survey of the American Board of Clinical Neuropsychology membership estimated that 39% of litigating or compensation-seeking MTBI cases resulted in diagnostic impressions of probable malingering (Mittenberg, Patton, Canyock, & Condit, 2002). Diminished symptom validity is also recognized by the Military TBI Task Force, who in their official position paper on the role of neuropsychology and rehabilitation psychology in the evaluation, management, and research of military veterans with traumatic brain injury stated, “Assessment methods utilized beyond the acute phase of injury could also include symptom validity or response bias measures, given the potential for motivational factors to sometimes confound the clinical picture of recovery after MTBI.” (McCrea et al., 2008, p. 17). In terms of motivational factors, it is notable that, in the veteran population, monetary gain in the form of monthly compensation is possible if the veteran is at least 10% disabled as a result of military service. Monetary gain is also possible in terms of a monthly pension if the veteran served during a time of war, has limited income, and is deemed “permanently and totally disabled” or is at least 65 years old (Veterans Benefits Administration, 2006). Traumatic Service Members’ Group Life Insurance (TSGLI) will also pay up to $100,000 for traumatic injury(s) from a single event, with potential payment exceeding $100,000 for multiple events (Traumatic Insurance Protection, 2007). If currently enlisted, secondary gain also may exist in the form of early discharge from military service.

In light of the aforementioned military-related motivational factors, evidence supporting the use of symptom validity instruments among the military population is in great demand. Recent publications are beginning to address the issue. For example, using a clinician-administered inventory designed to detect symptom exaggeration, the Structured Interview for Reported Symptoms (SIRS: Rogers, 1992), results of a recent study showed that the majority of Vietnam veterans (53%) presenting for treatment at a Veterans Affairs (VA) residential post-traumatic stress disorder (PTSD) treatment program exhibited clear signs of symptom exaggeration (Freeman, Powell, & Kimbrell, 2008). Importantly, in that study, patients’ scores on the SIRS correlated significantly with their scores on two other symptom validity tests, the Structured Inventory of Malingered Symptomatology (SIMS; Merckelbach & Smith, 2003) and the Miller Forensic Assessment of Symptoms Test (MFAST; Miller, 2004). In addition, although not specifically focusing on symptom validity, a recently published investigation of the cognitive sequelae of blast-related brain trauma among active-duty or veteran military personnel who were injured in Operation Iraqi Freedom/Operation Enduring Freedom reported that approximately 25% of patients were excluded from the study due to having been suspected of poor effort or malingering based on clinical presentation and/or failing certain measures of symptom validity, including the Word Memory Test (Green, 2003) at one site and the Medical Symptom Validity Test (MSVT; Green, 2004) and California Verbal Learning Test-II Long-Delay Forced Choice Recognition (Delis, Kramer, Kaplan, & Ober, 2000) at another site (Belanger, Kretzmer, Yoash-Gantz, Pickett, & Tupler, 2009). In order to further investigate the use of symptom validity tests among veterans of the current wars, the present study was designed to examine the clinical utility of one of the aforementioned symptom validity tests, the MSVT (Green, 2004), in a sample of soldiers returning from service in Operation Iraqi Freedom/Operation Enduring Freedom.

The MSVT is a brief computerized screening test for verbal memory impairment that includes built-in symptom validity testing. It contains two primary symptom validity subtests, a “consistency” subtest and two memory subtests. The interpretation of MSVT test results is typically completed in a two-step fashion. The first step involves the examination of three subtests, including the two primary symptom validity subtests and a third subtest that evaluates the consistency in the patient’s responses to the primary symptom validity subtests. A test-taker is considered to have failed the examination if they scored below specified cut-offs on any of these three indices. These subtests have been shown to be sensitive to diminished symptom validity, while typically remaining insensitive to true cognitive impairment, even when it is severe. Richman et al., (2006) demonstrated the simplicity of these tests by showing that the mean scores on these indices ranged between 95% and 99% correct in both adults and children tested in their own language and in adults and children tested in a foreign language, in which they were not fluent. It is notable that the sample employed in the aforementioned studies included adults with severe traumatic brain injury (TBI) or neurological disease and children with below-average IQs. Richman et al., (2006) also cite data compiled by Flaro showing that children with a full scale IQ of less than or equal to 70 scored a mean of 95% correct on the primary...
symptom validity subtests. In contrast, these authors demonstrated the sensitivity of the MSVT to diminished symptom validity by showing that 33% of patients with soft tissue injuries or fibromyalgia undergoing independent medical examination for a disability evaluation failed one or more of these indices. The overall pattern of results for individuals with soft tissue injuries or fibromyalgia who failed the MSVT were reported to be very similar to those of simulators, rather than to those of individuals with true neurological dysfunction.

The available data suggest that, of those groups assumed to be trying their best to do well on testing, only patients with advanced dementia score below the specified cut-off on the primary symptom validity subtests (Green, 2004). Anytime that a patient does score below these specified cut-offs, the examiner proceeds with the second step in test interpretation, which is profile analysis. This second step is employed because, even if a patient has failed the MSVT by performing below the specified cut-offs, their data may still be considered valid if their test scores conform to the Dementia Profile (DP). The DP is a series of rules that are designed to verify that the individual’s results are consistent with brain-behavior relationships by showing that they performed better on the easier versus the harder aspects of the MSVT. As explained by Howe and Loring (2009), the primary purpose of the DP is to reduce MSVT false positives by identifying individuals who failed the MSVT due to neurologic disease rather than diminished symptom validity. In their investigation of the DP among persons with dementia, Howe and Loring (2009) found that, of 20 persons with dementia who failed the MSVT, 17 met criteria for the DP, indicating that the DP demonstrated a sensitivity of 85%. Similarly, Frerichs (2006) showed that 37.5% of his clinical sample with mild cognitive impairment (N = 16) and 51.5% of his sample with dementia (N = 33) failed the MSVT; with the exception of one patient, all patients who failed the MSVT met criteria for the DP.

Although the literature suggests that the MSVT is a symptom validity test that shows promise for use in a variety of populations, it does not appear that its usefulness among the veteran population, specifically among generally younger returning troops of Operation Iraqi Freedom/Operation Enduring Freedom has been investigated. The aim of the current study was to preliminarily examine the rate of MSVT failure in the Operation Iraqi Freedom/Operation Enduring Freedom population. Given both the possibility for secondary gain and the potential for neurological dysfunction, it was hypothesized that MSVT test failure would exist among a sizeable minority of these soldiers. Furthermore, it was predicted that patient performance on other stand-alone and embedded symptom validity measures would provide concurrent validity data for the MSVT. Specifically, it was predicted that, upon examination, persons who failed the MSVT and did not satisfy criteria for the DP would show further evidence of non-credible test performance on other stand-alone and embedded symptom validity measures.

Materials and Methods

Participants

Data were collected from the files of 23 outpatients who had served in either Operation Iraqi Freedom (n = 19, 82.6%) or Operation Enduring Freedom (n = 4, 17.4%) and were referred to one of the authors for neuropsychological testing within a Polytrauma Network Site (Level 2) at a United States Department of Veterans Affairs (VA) Medical Center. The Polytrauma Network Site was not an acute care center, but one of 21 facilities in the country designed to provide long-term rehabilitative care to veterans and service members who experienced severe injuries (including brain injuries) to more than one organ system. Nine (39.1%) of the individuals in the current study remained enrolled in a branch of military service at the time of testing, whereas the majority had been discharged (n = 14, 60.9%). Referrals were primarily made from treatment providers within the Polytrauma Network Site itself, and were standard procedure for an individual reporting having been exposed to a blast and endorsing cognitive symptoms. The reviewed cases were consecutive referrals who were administered the MSVT as part of their clinical neuropsychological evaluation. The cases represent 23 of 29 total cases referred for neuropsychologist assessment. Six patients referred for testing were not administered the MSVT, primarily due to time constraints.

In terms of patient demographics, age of participants (N = 23) ranged from 23 to 54 years old, with a mean age of 33.00 years (SD = 7.97). Highest year of education completed by participants ranged from 12 to 18, with a mean level of 13.04 years (SD = 1.87). Nineteen (82.6%) participants were men and four (16.7%) were women. Of 23 participants, 18 (78.3%) were Caucasian and 5 participants (21.7%) were African-American. With the exception of one patient who reported having been injured by a fall, all other patients reported having been exposed to one or more blasts. The range of the number of blasts exposure events varied from 1 to 300, with one blast being the modal number reported by patients (n = 8, 34.8%). Months since injury ranged from 5 to 51, with 20 months being the modal number reported by patients (n = 3, 13.0%).
Measures and Procedures

The MSVT (Green, 2004) was administered as part of a larger battery of neuropsychological tests given to patients as a part of their standard clinical care. The MSVT was chosen for use due to (i) its brevity and (ii) the demonstrated high intercorrelation ($r = .70-.80$) between it and its longer and slightly modified counterpart, the Word Memory Test (Green, 2003), which is a well-established symptom validity measure. The MSVT requires 5 min of patient on-task time and a 10-min delay. Specific administration procedures are detailed in the test manual. To briefly summarize, after being presented twice with a series of 10 word pairs, the test-taker is asked to complete four trials, which result in five test scores. These trials include a free recall trial, a paired-associate recall trial, an immediate forced-choice recognition trial and a delayed forced-choice recognition trial. The aforementioned trials represent four test scores. The fifth test score is computed by examining the consistency in the test-takers responses to two of the trials. Some of these five scores are considered “easy” to do well on, whereas others are considered “hard” to do well on. In order to prevent misuse of sensitive test procedures, rather than revealing which specific test scores are easy vs. hard, the scores will be referred to simply as Easy 1 (E1), Easy 2 (E2), Easy 3 (E3), and Hard 1 (H1), Hard (H2). By keeping descriptions of test scores to a minimum, it is hoped that the security of the MSVT will be maintained.

Participants were considered to have failed the MSVT if their score on E1, E2, or E3 was below the specified cut-offs. Participants were considered to have met the criteria for the DP if (i) either E1, E2, or E3 were below the specified cut-offs; (ii) there were no scores below chance; (iii) there was a specified point difference between the mean of the easy items and the mean of the hard items (easy $>$ hard); (iv) E1 and E2 were greater than H2; and (v) H1 was greater than H2. The DP rules also require that clinical correlates of significant disability exist. The specific MSVT cut-offs and specified point difference mentioned earlier are detailed in the test manual (Green, 2004).

Although MSVT scores were the primary area of interest in the present investigation, an attempt was made to learn more about patients who failed the MSVT by gathering additional test data about them from their files. Additional test data summarized for persons who failed the MSVT included scores from the Mini Mental State Exam (MMSE; Folstein, Folstein, & McHugh, 1975); Test of Memory Malingering (Tombaugh, 1996); the long delay standard score on the Rey Auditory Verbal Learning Test (RAVLT; Schmidt, 1996); and the Ward Wechsler Adult Intelligence Scale – III (WAIS-III) Short Form Full Scale IQ (Pilgrim, Myers, Bayless, & Whetstone, 1999). To be specific, the following seven subtests of the WAIS-III were administered: Picture Completion, Digit Symbol, Block Design, Similarities, Arithmetic, Digit Span and Information.

Scores from multiple embedded symptom validity measures (i.e. cut-offs that are derived from pre-existing neuropsychological tests) were also extracted from the files of participants who failed the MSVT; these scores included Digit Symbol of the WAIS-III, Arithmetic of the WAIS-III, Reliable Digit Span and the Noncredible Score from the RAVLT. Utilizing published literature, we set cut-offs for failure on these measures that minimized false-positive errors. For Digit Symbol, age-scaled scores of $\leq 2$ were used to indicate test failure, as that cut-off is associated with 50% sensitivity among simulated malingerers and a 6% false-positive rate among persons with moderate to severe TBI (Etherton, Bianchini, Heinly, & Greve, 2006b). For Arithmetic subtest, age-scaled scores of $\leq 4$ were used to indicate test failure, as that cut-off is associated with 30% sensitivity among simulated malingerers and a 3% false-positive rate among persons with TBI (Etherton, Bianchini, Ciota, Heinly, & Greve, 2006a). Reliable Digit Span is calculated from the WAIS-III Digit Span subtest by summing the longest forward and backward digit strings. Both trials of the strings must be correct. For Reliable Digit Span, a cut-off score of $\leq 7$ was used to indicate test failure, as that cut-off has been associated with a 67% sensitivity and 93% specificity rate among persons with TBI (Mathias, Greve, Bianchini, Houston, & Crouch, 2002). The Noncredible Score from the RAVLT is calculated by adding True Recognition Score from List A (i.e. recognition minus false-positives) to number of items recognized in the first one-third of List A (Boone, Lu, & Wen, 2005). For the Noncredible Score, scores $\leq 12$ were used to indicate test failure, as that cut-off is associated with 73.8% sensitivity and 90% specificity among a mixed sample non-creditable patients, credible patients, and controls (Boone et al., 2005).

Statistical Analyses

Except where indicated, statistical analyses were calculated using SPSS, Version 10.1 (SPSS, Chicago, IL, USA). Alpha was set at $p < .05$. Participants were first classified into those who passed vs. failed the MSVT. Then, individual cases failing the MSVT were examined to determine whether or not they met the criteria for the DP. Because no patients who failed the MSVT met the criteria for the DP, using independent sample $t$-tests, scores on all subtests of the MSVT were compared between participants who failed the MSVT vs. those who passed. Following the latter analyses, additional test scores were descriptively summarized for persons who failed the MSVT. Then, in order to investigate how MSVT scores relate to other established embedded symptom validity measures, two-tailed Pearson correlations were conducted to examine the relationship between
Results

The MSVT profiles including the mean, standard deviation, and the range for all test scores are presented in Table 1. Upon examination, it was found that none of the individuals who failed the MSVT conformed to the DP. Results of independent t-tests comparing the group who passed (n = 19) vs. failed (n = 4) the MSVT showed that there were significant group differences for E1 [t(21) = 11.09, p < .001], E2 [t(21) = 7.92, p < .001], and E3 [t(21) = 9.63, p < .001], but not for H1 [t(21) = 1.09, p = .353] or H2 [t(21) = .51, p = .615].

Table 2 shows additional data gathered about the four individuals who failed the MSVT, all of whom were still enlisted in the military at the time of testing. The data showed that neuropsychological testing of these individuals occurred anywhere from 5 to 18 months post-injury. None of the patients reported post-traumatic amnesia (PTA) that exceeded their self-reported LOC, which ranged from 0 to 10 min. Based on the latter reported LOC and PTA, even the most severe brain injury reported by these patients, would be considered mild in nature. Three of the four patients who failed the MSVT also completed the TOMM; all three failed. Generally speaking, despite apparent diminished symptom validity, the patients’ scores on other select neuropsychological tests were all within the low-average range or above, including their scores on the 30-min delay of the RAVLT, which ranged from standard scores of 82–104 (M = 100, SD = 15); their scores on the short form of the WAIS-III, FSIQs, which ranged from 87 to 97 (M = 100, SD = 15); and their scores on the MMSE, which ranged from 24 to 29 (within normal limits >23).

Results of Pearson correlations between MSVT scores and embedded symptom validity measures showed that only the RAVLT Noncredible Score was significantly related to MSVT scores (Table 3). Table 4 shows that, among patients who failed the MSVT, other than the TOMM, the RAVLT Noncredible Score was the most often failed symptom validity test, showing a failure rate of 50%. Of the four patients who failed the MSVT, one patient failed three of five other symptom validity measures; one patient failed two of five other symptom validity measures; one patient failed one of five of the other symptom validity measures; and, one patient, who was not given the TOMM, failed one of the four other symptom validity measures.
Results of the current investigation showed that 17% of a sample of Operation Iraqi Freedom/Operation Enduring Freedom patients at a VA Polytrauma Network Site (Level 2) performed below cut-offs on the MSVT. The preliminary rate of 17% symptom validity test failure among Operation Iraqi Freedom/Operation Enduring Freedom soldiers falls between the 8% and 39% estimated rates of symptom exaggeration found among medical and MTBI cases, respectively, in litigating or compensation-seeking status who were seen by members of the American Board of Clinical Neuropsychology (Mittenberg et al., 2002). The finding of 17% MSVT failure among Operation Iraqi Freedom/Operation Enduring Freedom soldiers in the current study was nearly identical to the 18% rate of MSVT failure found among a group of individuals with financial incentive to under-perform who were referred to a memory disorder clinic (Howe, Anderson, Kaufman, Sachs, & Loring, 2007).

With regard to specific patterns of performance on the MSVT in the current study, on “easy” subtests of the MSVT, the group of individuals who failed the MSVT performed significantly worse than the group of individuals who passed. In contrast, there were no significant group differences on the “hard” subtests of the MSVT. The latter findings are likely not explainable by patterns of typical neurocognitive dysfunction, as individuals with cognitive impairment would be expected to show the opposite pattern. That is, it would be reasonable to expect individuals with true memory impairment to perform significantly worse than persons without cognitive impairment on the harder subtests of the MSVT, but not necessarily on the very simple subtests of the MSVT. The fact that, when compared with persons who passed the MSVT, persons who failed the MSVT in the current study showed relatively poorer performance on the easier vs. harder aspects of the task is suggestive of diminished symptom validity.

None of the patients who failed the MSVT in the current study met the criteria for the DP, suggesting that their poor scores on the MSVT were more reflective of diminished symptom validity than of true memory dysfunction. Additional test findings on other measures supported the conclusion that patients who failed the MSVT were exhibiting diminished symptom validity. For example, the four patients who failed the MSVT reported brief LOC, if any, and did not report PTA that exceeded 10 min, both of which suggest that their injuries did not approximate the severity that is typically seen in individuals who might exhibit symptom validity test failure due to brain damage. In addition, of the four patients who failed the MSVT, all failed at least one other stand-alone (i.e. TOMM) or embedded symptom validity measure (i.e. Digit Symbol, Arithmetic, Reliable Digit Span, and the RAVLT Noncredible Score). Among persons failing the MSVT, three took the TOMM and all three failed. Of the embedded symptom validity tests, the MSVT subscales showed significant and large correlations with only the RAVLT Noncredible Score. Fifty percent of the four patients who failed the MSVT also had scores lower than the cut-off on the RAVLT Noncredible Score. This is, perhaps, not surprising as both the MSVT and RAVLT involve forced choice tests of word recognition, whereas the other embedded symptom validity measures assessed different skill sets.

Taken as a whole, these preliminary data suggest that the specificity of the MSVT in the Operation Iraqi Freedom/Operation Enduring Freedom population is 100%, to that all persons who failed the MSVT demonstrated inconsistencies in their test performance.
results that suggested diminished symptom validity. As none of the patients who failed the MSVT performed at below chance levels, it cannot be said for certain that their scores represented “purposive distortion” (Reynolds, 1998, p. 272) consistent with definite malingering. However, both their pattern of performances on the MSVT and their performances on other stand-alone and embedded symptom validity measures clearly suggested that they were not putting forth adequate effort to do well on the neuropsychological testing battery. In fact, considering all aspects of their presentation, all patients who failed the MSVT met the diagnostic criteria for probable malingered neurocognitive dysfunction proposed by Slick, Sherman, and Iverson (1999).

Future research may address the various causes for symptom validity test underperformance. Among combat veterans, the elevated likelihood of anxiety in the form of PTSD may affect symptom presentation. Although evidence suggests that psychiatric symptoms such as depression and anxiety do not negatively affect performance on symptom validity tests for civilians (Ashendorf, Constantinou, & McCaffrey, 2004), it is possible that, among soldiers with PTSD, the fear of a potential return to combat may increase the odds of symptom magnification, resulting in diminished symptom validity. It is notable that all of the soldiers who failed the MSVT in the present study remained enlisted in a branch of the service at the time of testing; thus – for these soldiers – re-deployment was a distinct possibility.

The limitations of this study are clear. The sample size employed in the present study was rather small and only one stand-alone symptom validity test was administered to all patients. A future research program with a larger sample size and using more than one stand-alone symptom validity test is needed to provide additional insight into the actual prevalence of diminished symptom validity among this population. In this regard, a collaborative or cooperative study involving multiple VA treatment centers providing services to Operation Iraqi Freedom/Operation Enduring Freedom returnees may have the greatest power.

Conflict of Interest

None declared.

References


