Performance Validity and Symptom Validity Tests: Are They Measuring Different Constructs?

Anna S. Ord[1,2,3], **Robert D. Shura**[1,2,3,5], **Ashley R. Sansone**[1], **Sarah L. Martindale**[1,2,3], **Katherine H. Taber**[1,2,4,5], **Jared A. Rowland**[1,2,3]

[1]W. G. (Bill) Hefner VA Healthcare System, Salisbury, NC, USA

[2]Mid-Atlantic Mental Illness Research Education and Clinical Center, Durham, NC, USA

[3]Wake Forest School of Medicine, Winston-Salem, NC, USA

[4]Baylor College of Medicine, Houston, TX, USA

[5]Via College of Osteopathic Medicine, Blacksburg, VA, USA

Abstract

Objective: To evaluate the relationships among performance validity, symptom validity, symptom self-report, and objective cognitive testing.

Method: Combat Veterans ($N = 338$) completed a neurocognitive assessment battery and several self-report symptom measures assessing depression, PTSD symptoms, sleep quality, pain interference, and neurobehavioral complaints. All participants also completed two performance validity tests (PVTs) and one stand-alone symptom validity test (SVT) along with two embedded SVTs.

Results: Results of an exploratory factor analysis revealed a three-factor solution: Performance Validity, Cognitive Performance, and Symptom Report (SVTs loaded on the third factor). Results of *t*-tests demonstrated that participants who failed PVTs displayed significantly more severe

symptoms and significantly worse performance on most measures of neurocognitive functioning as compared to those who passed. Participants who failed a stand-alone SVT also reported significantly more severe symptomatology on all symptom report measures, but the pattern of cognitive performance differed based on the selected SVT cutoff. Multiple linear regressions revealed that both SVT and PVT failure explained unique variance in symptom report, but only PVT failure significantly predicted cognitive performance.

Conclusions: Performance and symptom validity tests measure distinct but related constructs. SVTs and PVTs are significantly related to both cognitive performance and symptom report; however, the relationship between symptom validity and symptom report is strongest. SVTs are also differentially related to cognitive performance and symptom report based on the utilized cutoff score.

Keywords

performance validity; symptom validity; factor analysis; cognition; Veterans

---

Neuropsychological assessment involves obtaining thorough information about examinee's functioning and relies on the assumption that examinees put forth adequate effort (Bush et al., 2014; Lezak et al., 2012; Millis, 2009). Moreover, validity of test results and subsequent conclusions depends on the accuracy and consistency of information provided by, or about, the examinee (Bush et al., 2014). Thus, validity assessment is an important part of standard practice within the field of clinical neuropsychology. Both the National Academy of Neuropsychology (Bush et al., 2005) and the American Academy of Clinical Neuropsychology (Heilbronner et al., 2009) recommend routine inclusion of validity tests in neuropsychological assessments. The term *validity* is broad and encompasses several subtypes, including symptom validity and performance validity. *Performance validity* refers to "the validity of actual ability task performance" (Larrabee, 2012, p. 626), whereas *symptom validity* refers to "the accuracy of symptomatic complaints on self-report measures" (Larrabee, 2012, p. 626).

In Veteran samples, a significant proportion of individuals fail performance and/or symptom validity measures. For example, Denning and Shura (2019) found a weighted mean average PVT failure rate of 30% across 50 studies of nearly 10,000 Veterans and service members: rates varied by context, with disability context showing the highest failure rates and research contexts showing the lowest failure rates. Regarding symptom validity, estimated failure rates on the MMPI-2-RF validity scales range from 5% to 27% (depending on the scale) as demonstrated in a study of over 17,000 protocols pulled from a national database of Veterans Affairs (VA) electronic medical records (Ingram et al., 2020). The VA maintains an extensive disability system for conditions incurred or aggravated during military service, which poses a challenge. Compensation and pension exams inform these decisions, and these forensic reports are part of the electronic medical record, which Veterans have full access to through an online portal. Therefore, disability becomes omnipresent, even in clinical or research contexts, and the lines across forensic versus clinical roles can blur, especially as perceived by patients. Thus, it is imperative to evolve the understanding of PVTs and SVTs in this population given this potential conflation of disability issues across contexts.

The relationship between performance validity and symptom validity has been examined in several studies utilizing factor analytic techniques (Nelson et al., 2007; Ruocco et al., 2008; Van Dyke et al., 2013). Specifically, Nelson et al. (2007) demonstrated that performance validity tests (PVTs; Victoria Symptom Validity Test [VSVT], Test of Memory Malingering [TOMM], and Letter Memory Test [LMT]) and symptom validity tests (SVTs; validity scales on the Minnesota Multiphasic Personality Inventory-2 [MMPI-2]) load on independent factors. Ruocco et al. (2008) also found that SVTs (validity indices on the Millon Clinical Multiaxial Inventory-III [MCMI-III]) and PVTs (Reliable Digit Span [RDS], TOMM) loaded on separate factors. Van Dyke et al. (2013) extended these findings by including symptom report and cognitive testing in the analysis. Their results indicated a 3-factor model: Cognitive Performance, Performance Validity, and Symptom Self-Report (with symptom validity measures loading on the last factor), indicating that PVTs and SVTs loaded on different factors (Van Dyke et al., 2013). These studies suggest that performance validity and symptom validity represent distinct constructs. However, cross-sectional studies have revealed significant (albeit weak to moderate) correlations between SVTs and PVTs (Copeland et al., 2016; Larrabee, 2003; Whitney et al., 2008), suggesting that these two constructs are not completely independent.

If PVTs and SVTs measured independent, mutually exclusive constructs, PVTs would be expected to be sensitive to cognitive performance but not symptom self-report, whereas SVTs would be expected to be sensitive to symptom report but not cognitive performance. However, this pattern has not been reliably observed in the published literature. As expected, PVTs are consistently associated with cognitive performance (Armistead-Jehle & Buican, 2012; Clark et al., 2014; Fox, 2011; Green et al., 2001; Lange et al., 2010; Lange et al., 2012; Meyers et al., 2011). However, PVTs have also been also consistently associated with clinical symptomatology, such as increased post-concussion and neurobehavioral symptoms (Clark et al., 2014; Lange et al., 2010), cognitive and neurological problems (Jones et al., 2012), sleep concerns (Johnson-Greene et al., 2013), pain complaints (Clark et al., 2014; Gervais et al., 2004; Johnson-Greene et al., 2013), somatic symptoms (Whiteside et al., 2010), symptoms of anxiety, depression, posttraumatic stress symptoms (Clark et al., 2014), and elevated emotional distress (Jones et al., 2012).

SVTs have also been linked with both symptom report and objective cognitive performance. For example, Gervais et al. (2008) found that the RBS scale on the MMPI-2 was associated with subjective memory complaints but not with an objective measure of verbal memory (Gervais et al., 2008), as would be expected if PVTs and SVTs were mutually exclusive constructs. On the other hand, Jurick and colleagues (2019) found that individuals with elevated MMPI-2-RF validity scales reported both a higher level of symptomatology and performed more poorly on objective cognitive measures. This finding was supported by Copeland et al. (2016) who reported that the FBS scale on the MMPI-2 was significantly correlated with greater self-reported neurobehavioral symptoms, as well as poorer performance on a verbal memory test. Martin et al. (2015) found that a significant amount of variance in RBS scores on the MMPI-2-RF was accounted for by cognitive test performance; however, after adjusting for PVT failure, the relationship between neurocognitive test performance and symptom validity scales on the MMPI-2-RF was no longer significant.

Overall, research findings suggest that SVTs and PVTs likely measure distinct, but not mutually exclusive, constructs. Further, the relationship between the two is complex. Several studies demonstrate consistent correlations between PVTs and SVTs, which suggests a measurable overlap between these constructs. However, studies examining differential relationships of PVTs and SVTs with symptom measures and cognitive performance are still rather limited and have produced contradictory results. Further examination is needed to further clarify the extent of convergence and divergence between symptom and performance validity with symptom report and cognitive performance. The primary objectives of the present study were to: (1) clarify the distinction between the constructs of symptom validity and performance validity by empirically examining the underlying factor structure within a comprehensive neuropsychological battery; and (2) to evaluate the differential associations of PVTs and SVTs with symptom-report measures and cognitive performance. It was hypothesized that SVT scores would be significantly associated with symptom report, but not with cognitive performance; similarly, it was hypothesized that PVT scores would be significantly associated with cognitive performance, but not with symptom-report measures.

Method

Participants

Data for the present analyses were collected as part of a larger study examining the effects of blast exposure on brain function, cognitive performance, and symptom presentation. Participants were primarily recruited from the local VA medical center through flyers, brochures, and mailings informing Veterans of the opportunity to participate. Flyers and brochures were also distributed to community Veteran centers. All study procedures were completed as part of the research study and did not overlap with clinical care. Participants were initially screened by telephone for inclusion and exclusion criteria. Participants were informed that their data could only be used as part of the research study and were not available for clinical or other evaluations. Eligibility criteria were: at least one combat deployment (combat defined as any score of > 17 on the Deployment Risk and Resiliency Inventory-2, Section D [Vogt et al., 2012]) after 9/11/2001, English speaking, 18 years of age or older, able to comply with instructions to complete study tasks, and able to provide informed consent. Exclusion criteria were: any penetrating head injury; non-deployment related TBI with loss of consciousness, and; presence of neurologic disorder, severe mental illness (Bipolar I and II, schizophrenia, other psychotic disorders), dementia, current substance use disorder, or psychotic symptoms. Exclusion criteria (e.g., neurologic disorder, severe mental illness) were initially evaluated within the medical record and subsequently during the telephone screen. TBI exclusion criteria were evaluated using the Mid-Atlantic MIRECC Assessment of TBI (MMA-TBI; Rowland et al., 2020). Psychiatric exclusion criteria were fully evaluated using the Structured Clinical Interview for *DSM-IV* Disorders (SCID-IV; First et al., 1996) as part of the study visit. All participants provided informed consent prior to participation. Study procedures were reviewed and approved by the local Institutional Review Board.

Measures

Symptom Measures—Participants completed several self-report questionnaires. The Posttraumatic Stress Disorder (PTSD) Checklist for DSM-5 (PCL-5; Blevins et al., 2015) is a 20-item questionnaire scored on a total scale of 0–80 that measures how bothered an individual is by PTSD symptoms over the past month. The Patient Reported Outcomes Measurement Information System Pain Interference (PROMIS-PI; Amtmann et al., 2010) is an 8-item questionnaire scored on a scale of 8–40 that measures the interference in daily activities caused by pain over the past seven days. The Pittsburgh Sleep Quality Index (PSQI; Buysse et al., 1989) is a 9-item questionnaire that provides a global sleep quality score ranging from 0 to 21. Sleep quality over the past month is evaluated. The Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001) is a 9-item self-report measure evaluating depressive symptoms over the past two weeks. The PHQ-9 is scored on a scale of 0–27. The Neurobehavioral Symptom Inventory (NSI; Cicerone & Kalmar, 1995) is a 22-item self-report assessment of how intensely post-concussive symptoms (both physical and behavioral) have bothered an individual over the past two weeks. The total score ranges from 0–88. Total scores were used for all measures. Higher scores indicate greater problems or poorer outcomes.

Cognitive Measures—Participants also completed a neuropsychological battery including the Wechsler Adult Intelligence Scale, fourth edition (WAIS-IV; Wechsler, 2008), Trail Making Test (TMT; Reitan & Wolfson, 1985) forms A and B, Controlled Oral Word Association Test (COWAT; Benton & Hamsher, 1989), and Semantic Fluency (Animal Naming; Benton & Hamsher, 1989). Scores on all cognitive measures were converted to demographically-corrected (sex, age, race, and education) T-scores ($M = 50$, $SD = 10$). WAIS-IV T-scores were derived from the WAIS-IV Advanced Clinical Solutions (ACS) demographically adjusted norms. T-scores for Animal Naming, COWAT, and TMT were derived from Heaton norms (Heaton et al., 2004).

Performance and Symptom Validity Measures—Performance validity measures were the Medical Symptom Validity Test (MSVT; Green, 2004) and the b Test (Boone et al., 2002). The MSVT is a learning and memory test designed to detect performance validity. Immediate Recall (IR), Delayed Recall (DR), and Consistency (CNS) scores were used to evaluate performance validity based on test manual criteria. The MSVT has shown adequate sensitivity and specificity in de82tecting invalid performance; for example, sensitivity predicting the WMT was .50 to .62 (Green, 2007). The b Test is a letter-recognition and discrimination test designed to measure performance validity (Boone et al., 2002). The b Test has been shown to adequately classify examinees based on their validity status in a study evaluating participants with depression, schizophrenia, head injury, stroke, and learning disabilities (Boone et al., 2002). Effort Index scores (E-scores) of the b Test were utilized in the present study, and the cutoff >81 was used to determine pass-fail status on the b Test; that score was associated with a sensitivity of .68 while maintaining specificity of .90 or higher (Roberson et al., 2013). The PVT fail group was identified as individuals failing either one of these measures (the MSVT or the b Test). Although there is extensive debate on how to identify invalid groups based on the use of multiple measures, given the eligibility

criteria of this study, failure of only one stand-alone PVT was used as the invalid criterion (Bush, 2005).

Symptom validity was measured by the Structured Inventory of Malingered Symptomatology (SIMS; Smith & Burger, 1997), the Validity-10 index embedded within the NSI (Vanderploeg et al., 2014), and the mild Brain Injury Atypical Symptom Scale (mBIAS; Cooper et al., 2011). The SIMS is a 75-item, true-or-false self-report measure. In simulation studies, SIMS scores discriminated well between participants instructed to simulate a certain condition (e.g., psychosis, an amnestic disorder, neurologic impairment, mania, depression, or low intelligence) and control subjects who were instructed to respond to questions honestly (Smith & Burger, 1997; van Impelen et al., 2014). Due to discrepant opinions regarding the sensitivity of the SIMS, we evaluated two published cutoffs for the SIMS: the liberal cutoff >14 indicated in the manual (Smith & Burger, 1997) and a more conservative cutoff >23 (Wisdom et al., 2010). Validity-10 includes 10 low-frequency items of the NSI (Vanderploeg et al., 2014) that are scored on a Likert scale from 0 (none) to 4 (very severe). Total scores range from 0 to 40, with higher scores indicating higher severity of endorsed symptoms. The cutoff >16 was used in the present study for the pass-fail status classification (Ashendorf, 2019). The mBIAS consists of five rationally-derived statements that are not commonly endorsed by individuals following a mild traumatic brain injury (Cooper et al., 2011). Severity of items is rated on a Likert scale from 1 (none) to 5 (very severe), resulting in a total score range of 5 to 25. Higher total score on the mBIAS indicates higher severity of endorsed symptoms. Cutoff >8 was utilized in the present study to determine pass-fail status on the mBIAS (Ashendorf, 2019).

Data Analysis

Analyses were conducted in SAS Enterprise Guide 7.1 (SAS Institute Inc., Cary, NC). Demographic variables were examined using univariate descriptive statistics. Aim 1 was evaluated using, a principal axis exploratory factor analysis (EFA) with oblique rotation (promax) to determine the underlying factor structure among the studied variables. Variables included in the EFA are listed in Table 2. The NSI was not included in the EFA because the Validity-10 embedded symptom validity scale was calculated from it. Collinearity between these variables would be artificially high, creating the potential for inaccurate factor loadings. The number of EFA factors to extract was determined using parallel analysis comparing Monte Carlo simulation results (program available at: http://edpsychassociates.com/Watkins3.html) from 100 iterations to initial eigenvalues. To evaluate Aim 2, differential relationships of SVTs and PVTs with symptom measures and cognitive tests, the following analyses were conducted: bivariate correlations, *t*-tests, and multiple regressions. Pearson product-moment correlations were utilized for all bivariate correlations between validity tests and measures of cognitive performance as well as self-reported symptomatology. Hypotheses comparing symptom presentation and cognitive performance were tested on the two groupings of interest for each validity test (pass or fail) using independent samples *t*-tests. Multiple linear regressions were conducted to evaluate contributions of PVTs and SVTs to cognitive performance and symptom report. SIMS cutoff of >23 was utilized for these regressions. This cutoff was considered more conservative and appropriate for assessing symptom validity in this sample. Almost half of the sample

failed the SIMS at the traditional cutoff >14, whereas only 13.6% of the sample failed it at the higher cutoff (>23), which is more consistent with validity failure rates in Veteran research samples (Ingram et al., 2020). A significance level of .05 was set *a priori* for all inferential tests. To reduce Type I error due to multiple comparisons, false discovery rate (FDR; Benjamini & Hochberg, 1995) was used to determine significant outcomes (step-down approach), correcting the FDR at $p < .05$ (number of comparisons was based on the number of outcome variables: 5 symptoms measures and 16 cognitive measures for each hypothesis).

Results

Sample characteristics are presented in Table 1. Eligible participants were 338 (86.39% male) Veterans between the ages of 23–71 ($M = 41.57$, $SD = 10.00$), with 9–22 years of education ($M = 14.99$, $SD = 2.16$). Approximately 20.41% of the sample failed at least one PVT ($n = 69$), 21.6% failed at least one SVT ($n = 73$), and 6.5% failed both – at least one PVT and at least one SVT ($n = 22$). Almost half of the sample ($n = 154$), failed the SIMS at the cutoff >14, whereas only 13.6% ($n = 46$) failed the SIMS at the cutoff >23. Given this high discrepancy in the failure rates, both cutoffs were explored in the present study. Over 85% of the sample had a service-connected disability (SCD). Among those who passed PVTs, 84% had SCD and among those who failed at least one PVT, 91% had SCD. This difference in proportion of service-connected Veterans between the two groups based on PVT pass-fail status was not statistically significant ($p = .123$). However, SVT pass-fail groups differed significantly in the rates of SCD. Specifically, among those who passed SVTs, 82.6% had SCD, compared to 95.9% reporting SCD among those who failed at least one SVT. This difference was statistically significant ($p = .004$).

Approximately half of the sample met criteria for a current psychiatric disorder, such as PTSD, major depressive disorder, dysthymia, and generalized anxiety disorder. Among those who failed at least one SVT, 78% had a current psychiatric diagnosis, which was significantly higher ($p < .001$) compared to only 46% who met criteria for a psychiatric disorder in the group that passed all SVTs. A similar pattern was observed based on PVT failure status: 74% of those who failed at least one PVT had a current psychiatric disorder as compared to 48% of those who passed all PVTs ($p < .001$). Similarly, participants who failed an SVT or a PVT had significantly higher ($p < .05$) rates of deployment TBI (66% and 62% respectively) as compared to those who passed all SVTs (46%) and all PVTs (47%). There were no statistically significant differences in terms of blast exposure based on SVT or PVT pass-fail status.

Aim 1: Exploratory Factor Analysis

All variables were entered into the EFA, including three subtests of the MSVT, the E-score of the b Test, four symptom measures, 14 cognitive performance measure scores, and total scores on the SIMS, mBIAS, and Validity-10 (Table 2). Using parallel analysis (Monte Carlo simulation) results, the EFA yielded a three-factor solution that accounted for 86.48% of variance. Factor 1 explained 35.23% of variance and included measures of cognitive performance. Factor 2 explained additional 32.20% of variance and included all symptom

measures and all SVTs. Factor 3 explained 19.05% of variance and included all performance validity measures. These factors were labeled as follows: Cognitive Performance (Factor 1), Symptom Report (Factor 2), and Performance Validity (Factor 3). The intercorrelations among the factors were weak to moderate: Factor 1 and Factor 2, $r = -0.21$; Factor 1 and Factor 3, $r = 0.32$; Factor 2 and Factor 3, $r = -0.35$. Table 2 displays factor loadings for all variables entered into the analysis.

Aim 2: Mean Comparisons

Independent samples t-tests evaluated differences in self-reported symptoms and objective cognitive performance between participants who passed and those who failed the PVTs and the SVTs. Results of these analyses are presented in Table 3. Two cutoffs were utilized for the SIMS: >14 and >23. Different patterns were observed for symptom measures and for cognitive tests at each SIMS cutoff. Significant differences were found on symptom measures between those who passed and those who failed the SIMS at both cutoffs, with comparable effect sizes. There were no significant differences on most cognitive performance measures (except for the Trail Making Test) based on the SIMS cutoff of >14, but at the higher cutoff >23, significant differences were observed on 11 of 16 cognitive tests. In short, participants who received very high scores on the SIMS were also displaying significantly lower performance across most neurocognitive measures. Similar results were observed for the MSVT: participants who failed the MSVT provided higher ratings on all symptom measures and performed significantly worse on 13 of 16 cognitive tests. Failure on the b Test was also associated with significant differences on four of five symptom measures and on 12 of 16 cognitive measures. These findings indicate that performance on both PVTs (the MSVT and the b Test) is significantly related not only to cognitive performance, but also to symptom report.

Aim 2: Multiple Linear Regressions

Two multiple regressions were conducted with Cognitive Performance and Symptom Report factors (derived from an EFA that included only measures of cognitive performance and symptom report) as dependent variables, and pass-fail status on any SVT and any PVT as independent variables. Results indicated that pass-fail status on the SVTs and the PVTs explained unique variance in symptom report; however, only performance validity failure was significantly related to cognitive performance, whereas symptom validity failure was not (Table 4). Interactions among the PVT and SVT variables were not significant in any model.

Discussion

The present study evaluated the relationship between the constructs of symptom and performance validity. Overall, results suggest that PVTs and SVTs measure separate constructs, but the distinction between SVTs and symptom report was not supported. Both PVTs and the SVTs were related to aspects of symptom report and cognitive function, indicating that both types of validity tests are sensitive to fluctuations in presentation outside of their respective domains (e.g., SVT failure was associated not only with symptom report,

but also with cognitive performance, whereas PVT failure was associated not only with cognition, but also with self-reported symptomatology).

First, the EFA clearly indicated that performance validity and symptom validity tests loaded on separate factors, supporting previous findings that PVTs and SVTs represent distinct constructs (Nelson et al., 2007; Ruocco et al., 2008). Thus, the present investigation replicated results of the very few studies published to date that empirically examined the theoretical distinction between these two constructs. The present study enhances extant literature by expanding the catalog of validity measures for which the effectiveness has been evaluated in a Veteran sample (e.g., SIMS and b Test). Additionally, the large sample utilized in the present study is comprised of combat-exposed Veterans, providing a more detailed investigation into a specific population in which assessing the role of validity is critical.

Although both the MSVT and the b Test loaded onto the same factor (performance validity), their loadings differed in magnitude. Specifically, the MSVT loadings were stronger than those of the b Test. Similarly, intercorrelations between MSVT subtests and the b Test were weak to moderate, with Pearson r coefficients ranging from −0.27 to −0.31 ($p < .001$), demonstrating a clear relationship between the variables, but far from complete overlap in variance. When these tests were used to classify participants into PVT pass-fail groups, results indicated that 43 participants (12.7%) failed the MSVT and 32 participants (9.5%) failed the b Test, but only six participants (1.8%) failed both tests. These findings highlight that although the MSVT and the b Test are parts of the same broader construct (performance validity, as evidenced by their loading on the same factor in the EFA), they likely measure different aspects of that construct. For example, the MSVT was designed to resemble a memory test, whereas the b Test was designed to resemble a test of attention or processing speed. Therefore, examinees who perform poorly on a test that they believe measures a certain cognitive ability (e.g., memory) may not necessarily fail a test that appears to measure a different ability (e.g., attention). Additionally, these tests may be differentially sensitive to other variables known to affect performance on PVTs. The MSVT and the b Test were administered at different time points during the assessment battery, which may also have contributed to differences in performance on these tests due to the potential influences of factors such as fatigue or exacerbation of psychiatric or physical symptoms (e.g., pain). It is outside the scope of this paper to speculate on various reasons why research participants may fail only one PVT and not both, but results suggest that performance validity is likely not a unidimensional construct. Future studies may focus on investigating this construct and relationships between various PVTs more closely. However, it is evident from the present study that the MSVT and the b Test likely provide unique information about examinee's performance validity and that different types of PVTs should be included in a neuropsychological assessment battery. Our results underscore the need for continuous and comprehensive assessment of performance validity that includes sampling of different cognitive abilities that may be associated with invalidity (Boone, 2009).

Next, SVT variables loaded onto the same factor with symptom report measures, indicating that the SVTs utilized in this study do not necessarily represent a construct that is distinct from symptom report. These results are consistent with Van Dyke et al. (2013) who also

found that SVTs loaded onto the same factor with symptom measures. When interpreting this finding, it is important to consider that symptom validity likely represents a complex, multi-dimensional construct encompassing non-content related responding (e.g., omission of items and inattentive or random responding) along with symptom underreporting or overreporting (Ben-Porath, 2012; Groth-Marnat, 2016; Merckelbach et al., 2019; Rogers & Bender, 2018). For example, several widely used and well-validated measures of psychopathology (such as the MMPI-2 and the Personality Assessment Inventory [PAI]) include multiple types of embedded SVTs assessing inconsistency and infrequency of responding along with impression management (Groth-Marnat, 2016). Therefore, symptom over-reporting represents only one facet of symptom validity, and it is possible that the items on the SVTs utilized in this study only evaluate this one aspect. Future research is needed to examine how different types of symptom validity are related to symptom report.

It is also plausible that different types of symptom validity may differentially relate to performance validity. Therefore, if a different aspect of symptom validity was assessed in this study (e.g., response consistency), a different pattern of results may have emerged. For example, McCaffrey et al. (2003) examined correlations between two PVTs (TOMM and Rey-15) and several validity scales on the MMPI-2. In their study PVT performance was significantly correlated with one validity scale assessing symptom exaggeration (Fb), but it was not significantly related to scales assessing consistency of responding (VRIN and TRIN) or symptom underreporting (K and L). Similarly, a study by Armistead-Jehle et al. (2012) demonstrated that pass-fail status on the Word Memory Test (a PVT) was significantly related to PAI validity scales assessing impression management (NIM and PIM) but not inconsistent responding (ICN) or endorsement of rare or bizarre statements (INF). Further research is needed to elucidate the relationships among different types of symptom validity, performance validity, and symptom report.

The present study also evaluated the differential associations of PVTs and SVTs with symptom measures and with cognitive performance. Both PVTs and SVTs were significantly associated with symptom report, although effect sizes indicated a stronger relationship between symptom presentation and SVTs, as compared to PVTs. Conversely, PVTs were significantly associated with most measures of cognitive performance, while associations between SVTs and cognitive tests were variable. In general, results are consistent with extant literature indicating associations among SVTs, PVTs, symptom measures, and cognitive tests (Armistead-Jehle & Buican, 2012; Copeland et al., 2016; Lange et al., 2010; Lange et al., 2012; Whiteside et al., 2010). However, the present study extended these results to demonstrate that PVTs and SVTs each explain unique variance in symptom report, whereas only PVTs explain unique variance in cognitive performance. Though there is a considerable amount of overlap between these two constructs, PVTs and SVTs appear to be unique in what they contribute to our understanding of examinee's presentation. Consequently, both types of validity tests should be included in comprehensive neuropsychological examinations to ensure that validity is adequately sampled throughout the assessment process (Boone, 2009).

The current results varied drastically based on the SVT cutoff selected (specifically, for the SIMS measure). First, failure rates were extremely elevated when using a cutoff >14

(over 45% of participants failed the SIMS based on this cutoff). Further, the relationship between the SIMS and cognitive testing fluctuated depending on the cutoff applied. At the more conservative >23 cutoff, SIMS displayed a pattern of results that was very similar to the MSVT, significantly differentiating performance on 11 of 16 cognitive tests. But at the traditional >14 cutoff, no differences were observed on 14 of 16 cognitive tests. Findings using the >23 cutoff are much more consistent with extant literature. Overall, our results suggest the traditional SIMS cutoff (>14) may not be appropriate for use in combat-exposed Veterans.

Several limitations should be considered when interpreting findings. First, the study was conducted in a sample of combat Veterans and results may not generalize to other military (e.g., active duty service members, Veterans who did not deploy) or civilian samples. Further research is needed to determine an optimal SIMS cutoff for combat-exposed Veterans as the >23 cutoff has not been empirically validated for this population. Another limitation of the present study involves the absence of memory tests in the cognitive battery. Because the MSVT was designed to be sensitive to invalidity on tasks of memory, future studies may include memory tests when assessing associations among different types of PVTs and various cognitive measures. More broadly, because no PVT is 100% sensitive, results could vary if PVTs other than the MSVT and the b Test were used. Finally, participants in the present sample were generally in the normal range of cognitive functioning, and examination of mean differences among groups did not reveal performance in the impaired range for any group. Future studies may involve more diverse samples of participants with a wider range of cognitive functioning.

Conclusions

In summary, findings indicate that performance validity and symptom validity are separate constructs: they are not mutually exclusive, but are also not strongly correlated, suggesting that both PVTs and SVTs contribute unique information about examinee's performance. SVTs and PVTs also differentially relate to symptom-report measures as compared to objective measures of cognitive functioning, with SVTs explaining more variance in symptom report. Different patterns of associations indicate that administration of only PVT(s) or only SVT(s) as part of an assessment battery is unlikely to be sufficient to capture invalid performance across both psychiatric presentation and cognitive performance. Therefore, assessment of both types of validity is needed in neuropsychological assessment in order to obtain an accurate and comprehensive understanding of examinee's functioning.

Acknowledgements

References

Amtmann D, Cook KF, Jensen MP, Chen WH, Choi S, Revicki D, Cella D, Rothrock N, Keefe F, Callahan L, & Lai JS (2010). Development of a PROMIS item bank to measure pain interference. Pain, 150(1), 173–182. 10.1016/j.pain.2010.04.025 [PubMed: 20554116]

Armistead-Jehle P, & Buican B (2012). Evaluation context and symptom validity test performances in a U.S. Military sample. Archives of Clinical Neuropsychology, 27(8), 828–839. 10.1093/arclin/acs086 [PubMed: 23047953]

Ashendorf L (2019). Neurobehavioral symptom validity in US Department of Veterans Affairs (VA) mild traumatic brain injury evaluations. Journal of Clinical and Experimental Neuropsychology, 41(4), 432–441. 10.1080/13803395.2019.1567693 [PubMed: 30712457]

Ben-Porath YS (2012). Interpreting the MMPI-2-RF. University of Minnesota Press.

Benjamini Y, & Hochberg Y (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B, 57(1), 289–300.

Benton AI, & Hamsher K (1989). Multilingual aphasia examination. AJA Associates.

Blevins CA, Weathers FW, Davis MT, Witte TK, & Domino JL (2015). The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation. Journal of Traumatic Stress, 28(6), 489–498. [PubMed: 26606250]

Boone KB (2009). The need for continuous and comprehensive sampling of effort/response bias during neuropsychological examinations. The Clinical Neuropsychologist, 23(4), 729–741. 10.1080/13854040802427803 [PubMed: 18949583]

Boone KB, Lu P, & Herzberg D (2002). The b Test. Western Psychological Services.

Bush SS, Heilbronner RL, & Ruff RM (2014). Psychological assessment of symptom and performance validity, response bias, and malingering: Official position of the Association for Scientific Advancement in Psychological Injury and Law. Psychological Injury and Law, 7(3), 197–205. 10.1007/s12207-014-9198-7

Bush SS, Ruff RM, Troster AI, Barth JT, Koffler SP, Pliskin NH, Reynolds CR, & Silver CH (2005). Symptom validity assessment: Practice issues and medical necessity. NAN policy & planning committee. Archives of Clinical Neuropsychology, 20(4), 419–426. 10.1016/j.acn.2005.02.002 [PubMed: 15896556]

Buysse DJ, Reynolds CF, Monk TH, Berman SR, & Kupfer DJ (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. Psychiatry Research, 28, 193–213. [PubMed: 2748771]

Carone DA (2009). Test review of the Medical Symptom Validity Test. Applied Neuropsychology, 16(4), 309–311. 10.1080/09084280903297883 [PubMed: 20183187]

Cicerone KD, & Kalmar K (1995). Persistent postconcussion syndrome: The structure of subjective complaints after mild traumatic brain injury. The Journal of Head Trauma Rehabilitation, 10(3), 1–17. 10.1097/00001199-199510030-00002

Clark AL, Amick MM, Fortier C, Milberg WP, & McGlinchey RE (2014). Poor performance validity predicts clinical characteristics and cognitive test performance of OEF/OIF/OND Veterans in a research setting. The Clinical Neuropsychologist, 28(5), 802–825. 10.1080/13854046.2014.904928 [PubMed: 24766521]

Cooper DB, Nelson L, Armistead-Jehle P, & Bowles AO (2011). Utility of the mild brain injury atypical symptoms scale as a screening measure for symptom over-reporting in operation enduring freedom/operation iraqi freedom service members with post-concussive complaints. Archives of Clinical Neuropsychology, 26(8), 718–727. 10.1093/arclin/acr070 [PubMed: 21873326]

Copeland CT, Mahoney JJ, Block CK, Linck JF, Pastorek NJ, Miller BI, Romesser JM, & Sim AH (2016). Relative utility of performance and symptom validity tests. Archives of Clinical Neuropsychology, 31(1), 18–22. 10.1093/arclin/acv065 [PubMed: 26537776]

Denning JH, & Shura RD (2019). The cost of malingering mild traumatic brain injury-related cognitive deficits during compensation and pension evaluations in the veterans benefits administration. Applied Neuropsychology: Adult, 26, 1–16. doi: 10.1080/23279095.2017.1350684 [PubMed: 28816502]

First MB, Spitzer RL, Gibbon M, & Williams JBW (1996). Structured Clinical Interview for the DSM-IV-TR Axis I Disorders. American Psychiatric Press, Inc.

Fox DD (2011). Symptom validity test failure indicates invalidity of neuropsychological tests. The Clinical Neuropsychologist, 25(3), 488–495. 10.1080/13854046.2011.554443 [PubMed: 21491351]

Gervais RO, Ben-Porath YS, Wygant DB, & Green P (2008). Differential sensitivity of the Response Bias Scale (RBS) and MMPI-2 validity scales to memory complaints. The Clinical Neuropsychologist, 22(6), 1061–1079. 10.1080/13854040701756930 [PubMed: 18609315]

Gervais RO, Rohling ML, Green P, & Ford W (2004). A comparison of WMT, CARB, and TOMM failure rates in non-head injury disability claimants. Archives of Clinical Neuropsychology, 19(4), 475–487. 10.1016/j.acn.2003.05.001 [PubMed: 15163449]

Green P (2004). Test manual for the Medical Symptom Validity Test. Green's Publishing.

Green P, Rohling ML, Lees-Haley PR, & Allen LM (2001). Effort has a greater effect on test scores than severe brain injury in compensation claimants. Brain Injury, 15(12), 1045–1060. 10.1080/02699050110088254 [PubMed: 11712951]

Green P (2007). Spoiled for choice: Making comparisons between forced-choice effort tests. In Boone KB (Ed.), Assessment of feigned cognitive impairment: A neuropsychological perspective (pp. 50–77). The Guilford Press.

Groth-Marnat G (2016). Handbook of psychological assessment (6th ed.). John Wiley & Sons.

Heaton RK, Miller SW, Taylor MJ, & Grant I (2004). Revised comprehensive norms for an expanded Halstead-Reitan battery: Demographically adjusted neuropsychological norms for African American and Caucasian adults. Psychological Assessment Resources.

Heilbronner RL, Sweet JJ, Morgan JE, Larrabee GJ, & Millis SR (2009). American Academy of Clinical Neuropsychology consensus conference statement on the neuropsychological assessment of effort, response bias, and malingering. The Clinical Neuropsychologist, 23(7), 1093–1129. 10.1080/13854040903155063 [PubMed: 19735055]

Ingram PB, Tarescavage AM, Ben-Porath YS, & Oehlert ME (2020). Patterns of MMPI-2-Restructured Form (MMPI-2-RF) validity scale scores observed across Veteran Affairs settings. Psychological Services, 17(3), 355–362. [PubMed: 30816738]

Johnson-Greene D, Brooks L, & Ference T (2013). Relationship between performance validity testing, disability status, and somatic complaints in patients with fibromyalgia. The Clinical Neuropsychologist, 27(1), 148–158. 10.1080/13854046.2012.733732 [PubMed: 23121595]

Jones A, Ingram MV, & Ben-Porath YS (2012). Scores on the MMPI-2-RF scales as a function of increasing levels of failure on cognitive symptom validity tests in a military sample. The Clinical Neuropsychologist, 26(5), 790–815. 10.1080/13854046.2012.693202 [PubMed: 22709365]

Jurick SM, Crocker LD, Keller AV, Hoffman SN, Bomyea J, Jacobson MW, & Jak AJ (2019). The Minnesota Multiphasic Personality Inventory-2-RF in treatment-seeking veterans with history of mild traumatic brain injury. Archives of Clinical Neuropsychology, 34(3), 366–380. 10.1093/arclin/acy048

Kroenke K, Spitzer RL, & Williams JB (2001). The PHQ-9: validity of a brief depression severity measure. Journal of General Internal Medicine, 16(9), 606–613. http://www.ncbi.nlm.nih.gov/pubmed/11556941 [PubMed: 11556941]

Lange RT, Iverson GL, Brooks BL, & Rennison VLA (2010). Influence of poor effort on self-reported symptoms and neurocognitive test performance following mild traumatic brain injury. Journal of Clinical and Experimental Neuropsychology, 32(9), 961–972. 10.1080/13803391003645657 [PubMed: 20437284]

Lange RT, Pancholi S, Bhagwat A, Anderson-Barnes V, & French LM (2012). Influence of poor effort on neuropsychological test performance in U.S. military personnel following mild traumatic brain injury. Journal of Clinical and Experimental Neuropsychology, 34(5), 453–466. 10.1080/13803395.2011.648175 [PubMed: 22273465]

Larrabee GJ (2003). Exaggerated pain report in litigants with malingered neurocognitive dysfunction. The Clinical Neuropsychologist, 17(3), 395–401. [PubMed: 14704890]

Larrabee GJ (2012). Performance validity and symptom validity in neuropsychological assessment. Journal of the International Neuropsychological Society, 18(4), 625–630. [PubMed: 23057079]

Lezak MD, Howieson DB, Bigler ED, & Tranel D (2012). Neuropsychological assessment (5th ed.). Oxford University Press.

Martin PK, Schroeder RW, Heinrichs RJ, & Baade LE (2015). Does true neurocognitive dysfunction contribute to Minnesota Multiphasic Personality Inventory-Restructured Form cognitive validity scale scores? Archives of Clinical Neuropsychology, 30(5), 377–386. [PubMed: 26051870]

McCaffrey RJ, O'Bryant SE, Ashendorf L, & Fisher JM (2003). Correlations among the TOMM, Rey-15, and MMPI-2 validity scales in a sample of TBI litigants. Journal of Forensic Neuropsychology, 3(3), 45–53.

Merckelbach H, Dandachi-FitzGerald B, van Helvoort D, Jelicic M, & Otgaar H (2019). When patients overreport symptoms: More than just malingering. Current Directions in Psychological Science, 28(3), 321–326. 10.1177/0963721419837681

Meyers JE, Volbrecht M, Axelrod BN, & Reinsch-Boothby L (2011). Embedded symptom validity tests and overall neuropsychological test performance. Archives of Clinical Neuropsychology, 26(1), 8–15. 10.1093/arclin/acq083 [PubMed: 21242283]

Millis SR (2009). Methodological challenges in assessment of cognition following mild head injury: response to Malojcic et al. 2008. Journal of Neurotrauma, 26(12), 2409–2410. 10.1089/neu.2008.0530 [PubMed: 19659408]

Nelson NW, Sweet JJ, Berry DT, Bryant FB, & Granacher RP (2007). Response validity in forensic neuropsychology: Exploratory factor analytic evidence of distinct cognitive and psychological constructs. Journal of the International Neuropsychological Society, 13(3), 440–449. [PubMed: 17445293]

Reitan RM, & Wolfson D (1985). The Halstead-Reitan neuropsychological test battery: Theory and clinical interpretation. Neuropsychology Press.

Roberson CJ, Boone KB, Goldberg H, Miora D, Cottingham M, Victor T, Ziegler E, Zeller M, & Wright M (2013). Cross validation of the b Test in a large known groups sample. The Clinical neuropsychologist, 27(3), 495–508. [PubMed: 23157695]

Rogers R, & Bender SD (Eds.). (2018). Clinical assessment of malingering and deception (4th ed.). The Guilford Press.

Rowland JA, Martindale SL, Shura RD, Miskey HM, Bateman JR, Epstein EL, …& Taber KH (2020). Initial validation of the Mid-Atlantic MIRECC Assessment of Traumatic Brain Injury. Journal of Neurotrauma. Advance online publication. 10.1089/neu.2019.6972

Ruocco AC, Swirsky-Sacchetti T, Chute DL, Mandel S, Platek SM, & Zillmer EA (2008). Distinguishing between neuropsychological malingering and exaggerated psychiatric symptoms in a neuropsychological setting. The Clinical Neuropsychologist, 22(3), 547–564. 10.1080/13854040701336444 [PubMed: 17853126]

Smith GP, & Burger GK (1997). Detection of malingering: validation of the Structured Inventory of Malingered Symptomatology (SIMS). Journal of the American Academy of Psychiatry and the Law, 25(2), 183–189.

Van Dyke SA, Millis SR, Axelrod BN, & Hanks RA (2013). Assessing effort: differentiating performance and symptom validity. The Clinical Neuropsychologist, 27(8), 1234–1246. 10.1080/13854046.2013.835447 [PubMed: 24028487]

van Impelen A, Merckelbach H, Jelicic M, & Merten T (2014). The Structured Inventory of Malingered Symptomatology (SIMS): A systematic review and meta-analysis. The Clinical Neuropsychologist, 28(8), 1336–1365. 10.1080/13854046.2014.984763 [PubMed: 25494444]

Vogt D, Smith B, King D, & King L (2012). Manual for the Deployment Risk and Resilience Inventory-2 (DRRI-2): A collection of measures for studying deployment-related experiences of military veterans. National Center for PTSD.

Wechsler D (2008). Wechseler Adult Intelligence Scale - Fourth Edition. Pearson.

Vanderploeg RD, Cooper DB, Belanger HG, Donnell AJ, Kennedy JE, Hopewell CA, & Scott SG (2014). Screening for postdeployment conditions: Development and cross-validation of an embedded validity scale in the Neurobehavioral Symptom Inventory. The Journal of Head Trauma Rehabilitation, 29(1), 1–10. 10.1097/HTR.0b013e318281966e [PubMed: 23474880]

Whiteside DM, Clinton C, Diamonti C, Stroemel J, White C, Zimberoff A, & Waters D (2010). Relationship between suboptimal cognitive effort and the clinical scales of

the Personality Assessment Inventory. The Clinical Neuropsychologist, 24(2), 315–325. 10.1080/13854040903482822 [PubMed: 20077355]

Whitney KA, Davis JJ, Shepard PH, & Herman SM (2008). Utility of the Response Bias Scale (RBS) and other MMPI-2 validity scales in predicting TOMM performance. Archives of Clinical Neuropsychology, 23(7–8), 777–786. [PubMed: 18930375]

Wisdom NM, Callahan JL, & Shaw TG (2010). Diagnostic utility of the Structured Inventory of Malingered Symptomatology to detect malingering in a forensic sample. Archives of Clinical Neuropsychology, 25(2), 118–125. 10.1093/arclin/acp110 [PubMed: 20110279]

Key Points

Question:

Are performance validity tests and symptom validity tests measuring different constructs, and are they differentially related to cognition and clinical symptoms?

Findings:

Symptom validity and performance validity are distinct but related constructs, and both are associated with cognitive performance and with symptom self-report.

Importance:

A comprehensive neuropsychological assessment battery should include both symptom validity and performance validity tests because they provide unique information about examinee's performance.

Next Steps:

Further evaluation of specific symptom validity measures is necessary to assess distinct differences between symptom validity and symptom self-report, as well as to determine whether higher cutoffs would be more appropriate for various populations, such as combat-exposed Veterans.

Table 1

Sample Characteristics (N = 338)

Characteristic	M or n	SD or %	Range
Age	41.57	10.00	23 – 71
Years Education	14.99	2.16	9 – 22
Sex			
Male	292	86.39%	—
Female	46	13.61%	—
Race/Ethnicity*			
White	193	57.10%	—
Black	136	40.24%	—
Other	21	6.21%	—
Hispanic/Latinx	16	4.73%	—
Branch of Service			
Air Force	32	9.47%	—
Army	245	72.49%	—
Marine Corps	37	10.95%	—
Navy	24	7.10%	—
Service-Connected Disability	288	85.20%	—
Current Psychiatric Disorder*	179	52.96%	—
PTSD	126	37.28%	—
Major Depressive Disorder	46	13.61%	—
Dysthymia	15	4.44%	—
Generalized Anxiety Disorder	22	6.51%	—
History of Deployment TBI	170	50.30%	—
Mild Deployment TBI	150	88.24%	—
Moderate Deployment TBI	20	11.76%	—
History of Blast Exposure	244	72.19%	—
History of Blast TBI	119	48.77%	—
Number of Combat Tours Served	2.69	3.28	1 – 50
Time Since Last Combat Deployment (Days)	3627	1263	447 – 6096
DRRI-2	36.48	16.56	10 – 93
NSI	26.86	17.64	0 – 78
PCL-5	33.42	20.03	0 – 78
PHQ-9	12.14	7.12	0 – 30
PROMIS-PI	19.96	9.81	8 – 40
PSQI	11.14	4.37	1 – 21
Semantic Fluency	50.47	11.09	12 – 86
Phonemic Fluency	47.91	10.73	26 – 86
TMT-A	47.39	11.35	7 – 86
TMT-B	48.15	11.13	5 – 81
FSIQ	99.43	13.33	65 – 142

Characteristic	*M* or *n*	*SD* or %	Range
VCI	102.12	13.56	63 – 150
PRI	100.10	14.68	67 – 138
WMI	97.19	13.25	66 – 145
PSI	97.37	13.74	53 – 137
Fail MSVT	43	12.72%	—
Fail b Test	32	9.47%	—
Fail any PVT	69	20.41%	—
Fail SIMS (> 14 cutoff)	154	45.56%	—
Fail SIMS (> 23 cutoff)	46	13.6%	—
Fail mBIAS (> 8 cutoff)	13	3.8%	—
Fail Validity-10 (> 16 cutoff)	49	14.5%	—
Fail any SVT	73	21.60%	

Note.

*
Categories are not mutually exclusive. Other race/ethnicity category includes Asian, Native American/Alaska Native, Native Hawaiian/Pacific Islander, Other, Not Sure, and Refused. These categories have been collapsed due to a small number of participants in each group. Branch of Service indicates most recent branch of service. Branches of service have been collapsed to include Reserve and Guard units. Coast Guard was not represented in the present sample. PTSD = post-traumatic stress disorder; TBI = traumatic brain injury; DRRI-2 = Deployment Risk and Resiliency Inventory-2, Section D (Combat Experiences), total score; NSI = Neurobehavioral Symptom Inventory, total score; PCL-5 = PTSD Checklist-5, total score; PHQ-9 = Patient Health Questionnaire-9, total score; PROMIS-PI = Patient Reported Outcomes Measurement Information System-Pain Interference, total score; PSQI = The Pittsburgh Sleep Quality Index, global score; Semantic Fluency = Animal Naming Test, T score; Phonemic Fluency = Controlled Oral Word Association Test, T score; TMT-A = Trail Making Test A, T score; TMT-B = Trail Making Test B, T score; FSIQ = Full Scale Intelligence Quotient, standard score; VCI = Verbal Comprehension Index, standard score; PRI = Perceptual Reasoning Index, standard score; WMI = Working Memory Index, standard score; PSI = Processing Speed Index, standard score; MSVT = Medical Symptom Validity Test; PVT = performance validity test; SIMS = Structured Inventory of Malingered Symptomatology; mBIAS = mild brain injury atypical symptoms scale; Validity-10 = Validity-10 scale embedded within NSI; SVT = symptom validity test.

Table 2

Factor Loadings for All Variables Included in the Principal Axis Factor Analysis (N = 338)

Measures	Factor 1	Factor 2	Factor 3
Performance Validity Measures			
MSVT IR	0.20	−0.18	**0.78**
MSVT DR	0.14	−0.20	**0.93**
MSVT CNS	0.13	−0.23	**0.86**
b Test	−0.15	0.13	**−0.28**
Symptom Validity Measures			
SIMS	−0.13	**0.79**	−0.27
mBIAS	−0.07	**0.45**	−0.19
Validity-10	−0.06	**0.85**	−0.11
Self-Report Symptom Measures			
PCL-5	−0.12	**0.89**	−0.08
PHQ-9	−0.05	**0.89**	−0.06
PROMIS-PI	−0.06	**0.69**	−0.13
PSQI	−0.08	**0.73**	−0.08
Cognitive Performance Measures			
Phonemic Fluency	**0.52**	−0.03	0.13
TMT-B	**0.57**	−0.08	0.13
WAIS-IV:			
Block Design	**0.60**	−0.09	0.01
Similarities	**0.53**	−0.07	0.12
Digit Span	**0.65**	−0.05	0.10
Matrix Reasoning	**0.63**	0.02	0.07
Vocabulary	**0.59**	−0.11	0.15
Arithmetic	**0.61**	−0.06	0.08
Symbol Search	**0.51**	−0.13	0.07
Visual Puzzles	**0.56**	−0.07	−0.01
Information	**0.52**	0.01	0.06
Coding	**0.55**	−0.10	0.09
Cancellation	**0.56**	−0.06	0.10
Letter-Number Sequencing	**0.61**	−0.01	0.07

Note. MSVT = Medical Symptom Validity Test; IR = immediate recognition; DR = delayed recognition; CNS = consistency score; PA = paired associates; FR = free recall; b Test = effort index score on the b Test; SIMS = Structured Inventory of Malingered Symptomatology, total score; mBIAS = mild brain injury atypical symptoms scale, total score; Validity-10 = Validity-10 scale embedded within NSI, total score; PCL-5 = PTSD Checklist-5; PHQ-9 = Patient Health Questionnaire-9; PROMIS-PI = Patient Reported Outcomes Measurement Information System-Pain Interference; PSQI = The Pittsburgh Sleep Quality Index; Phonemic Fluency = Controlled Oral Word Association Test; TMT-B = Trail Making Test B; WAIS-IV = Wechsler Adult Intelligence Scale 4th edition; bold font and gray highlighting indicate strongest loading for a specific factor.

VA Author Manuscript

VA Author Manuscript

VA Author Manuscript

Table 3

T-Tests Comparing Pass and Fail Groups on Symptom and Performance Validity Tests (N = 338)

Outcome Measures	SIMS > 23 cutoff			SIMS > 14 cutoff			MSVT			b Test		
	Pass (n = 292) M (SD)	Fail (n = 46) M (SD)	d	Pass (n = 184) M (SD)	Fail (n = 154) M (SD)	d	Pass (n = 295) M (SD)	Fail (n = 43) M (SD)	d	Pass (n = 306) M (SD)	Fail (n = 32) M (SD)	d
Self-Report Symptom Measures												
NSI	23.33 (15.50)	49.26 (13.54)	1.78	15.78 (11.29)	40.10 (14.45)	1.88	25.36 (17.15)	37.14 (17.70)	0.68	26.24 (17.43)	32.78 (18.78)	0.36
PCL-5	29.74 (18.35)	56.80 (13.44)	1.68	21.63 (15.38)	47.52 (15.28)	1.69	31.63 (19.67)	45.72 (18.29)	0.74	32.43 (19.87)	42.91 (19.31)	0.53
PHQ-9	10.87 (6.55)	20.22 (5.00)	1.60	8.11 (5.43)	16.96 (5.77)	1.58	11.56 (6.95)	16.16 (7.01)	0.66	11.84 (7.04)	15.03 (7.28)	0.45
PROMIS-PI	18.38 (9.28)	30.02 (6.63)	1.44	15.49 (7.77)	25.31 (9.33)	1.14	19.11 (9.50)	25.84 (10.01)	0.69	19.86 (9.77)	20.94 (10.35)	0.11
PSQI	10.54 (4.22)	14.94 (3.35)	1.15	9.08 (3.86)	13.60 (3.62)	1.21	10.78 (4.35)	13.61 (3.75)	0.70	10.98 (4.44)	12.63 (3.44)	0.31
Cognitive Performance Measures												
Semantic Fluency	51.23 (10.65)	46.33 (12.35)	−0.42	51.00 (11.03)	50.03 (10.99)	−0.09	51.02 (10.71)	46.67 (12.93)	−0.37	50.39 (10.92)	51.22 (12.79)	0.07
Phonemic Fluency	48.52 (10.52)	44.02 (11.34)	−0.41	48.22 (10.20)	47.53 (11.34)	−0.06	48.51 (10.53)	43.79 (11.26)	−0.43	47.65 (10.29)	50.34 (14.21)	0.22
TMT-A	48.38 (10.91)	41.11 (12.15)	−0.63	49.15 (10.41)	45.29 (12.08)	−0.34	48.48 (10.74)	39.88 (12.66)	−0.73	48.00 (11.21)	41.53 (11.21)	−0.58
TMT-B	48.75 (10.83)	44.33 (12.29)	−0.38	49.61 (10.96)	46.42 (11.12)	−0.29	48.77 (10.82)	43.88 (12.34)	−0.42	48.66 (10.60)	43.25 (14.58)	−0.42
WAIS-IV:												
Block Design	49.01 (9.44)	44.28 (10.12)	−0.48	48.84 (9.08)	47.82 (10.31)	−0.10	48.87 (9.57)	44.91 (9.65)	−0.41	48.88 (9.49)	43.47 (10.00)	−0.56

Outcome Measures	SIMS > 23 cutoff			SIMS > 14 cutoff			MSVT			b Test		
	Pass (n = 292) M (SD)	Fail (n = 46) M (SD)	d	Pass (n = 184) M (SD)	Fail (n = 154) M (SD)	d	Pass (n = 295) M (SD)	Fail (n = 43) M (SD)	d	Pass (n = 306) M (SD)	Fail (n = 32) M (SD)	d
Similarities	48.86 (10.22)	44.50 (11.11)	−0.41	49.11 (10.65)	47.27 (10.12)	−0.18	49.13 (10.22)	42.37 (10.10)	−0.67	48.41 (10.25)	46.91 (12.16)	−0.13
Digit Span	46.59 (9.91)	42.24 (9.50)	−0.45	46.88 (9.91)	44.96 (9.94)	−0.19	46.58 (10.03)	42.05 (8.51)	−0.49	46.54 (9.95)	40.84 (8.56)	−0.61
Matrix Reasoning	49.27 (10.89)	47.28 (8.96)	−0.20	48.50 (10.70)	49.59 (10.62)	0.10	49.69 (10.50)	44.28 (10.65)	−0.51	49.46 (10.56)	44.56 (10.75)	−0.46
Vocabulary	49.32 (9.52)	44.43 (10.87)	−0.48	49.69 (10.34)	47.42 (9.08)	−0.23	49.63 (9.49)	42.00 (9.71)	−0.79	49.11 (9.63)	44.34 (10.92)	−0.46
Arithmetic	46.18 (10.02)	40.89 (10.17)	−0.52	46.21 (10.03)	44.57 (10.34)	−0.16	45.82 (10.23)	43.05 (9.66)	−0.28	45.95 (10.15)	40.84 (9.58)	−0.52
Symbol Search	51.31 (10.69)	48.44 (10.19)	−0.27	51.97 (10.68)	49.66 (10.51)	−0.22	51.29 (10.46)	48.33 (11.69)	−0.27	51.47 (10.37)	45.59 (11.93)	−0.53
Visual Puzzles	51.00 (10.41)	46.50 (8.52)	−0.47	51.03 (10.25)	49.62 (10.29)	−0.14	50.57 (10.23)	49.14 (10.65)	−0.14	50.84 (10.22)	46.09 (9.95)	−0.47
Information	48.52 (9.99)	47.44 (11.18)	−0.10	48.42 (10.94)	48.32 (9.15)	−0.01	48.88 (10.09)	44.88 (9.95)	0.40	48.65 (10.22)	45.75 (9.14)	−0.30
Coding	47.75 (10.77)	43.61 (9.68)	−0.40	47.78 (10.51)	46.47 (10.93)	−0.12	47.82 (10.03)	42.81 (13.89)	−0.41	47.64 (10.41)	42.84 (12.61)	−0.41
Letter-Number Sequencing	47.15 (9.62)	45.20 (8.68)	−0.21	46.91 (9.62)	46.85 (9.41)	−0.01	47.48 (9.29)	42.79 (10.09)	−0.48	47.36 (9.33)	42.41 (10.21)	−0.51
Cancellation	45.90 (9.69)	43.30 (10.02)	−0.26	46.09 (9.57)	44.89 (9.98)	−0.12	46.12 (9.45)	41.61 (11.02)	−0.44	46.03 (9.58)	40.91 (10.44)	−0.51

Note. SIMS = Structured Inventory of Malingered Symptomatology; d = Cohen's d; NSI = Neurobehavioral Symptom Inventory, total score; PCL-5 = PTSD Checklist-5, total score; PHQ-9 = Patient Health Questionnaire-9, total score; PROMIS-PI = Patient Reported Outcomes Measurement Information System-Pain Interference, total score; PSQI = The Pittsburgh Sleep Quality Index, global score; Semantic Fluency = Animal Naming Test, T score; Phonemic Fluency = Controlled Oral Word Association Test, T score; TMT-A = Trail Making Test A, T score; TMT-B = Trail Making Test B, T score; WAIS-IV = Wechsler Adult Intelligence Scale 4th edition; T scores for all subtests of the WAIS-IV are presented; bold font and gray highlighting indicate statistical significance after adjustment for familywise error using False Discovery Rate (FDR).

Table 4

Multiple Regressions Evaluating Contributions of PVT and SVT Failure to Cognitive Performance and Symptom Report (N = 338)

Independent Variable	Omnibus Model			Parameter Estimates						
	F	p	R^2	B	SEB	β	t	p	LLCI	ULCI
Cognitive Performance	10.54	<.001	.088							
Symptom Validity				−0.21	0.14	−0.09	−1.45	.147	−0.49	0.07
Performance Validity				**−0.54**	**0.15**	**−0.23**	**−3.70**	**<.001**	**−0.83**	**−0.25**
Symptom Validity * Performance Validity				−0.20	0.27	−0.05	−0.73	.464	−0.74	0.33
Symptom Report	84.30	<.001	.435							
Symptom Validity				**1.49**	**0.12**	**0.63**	**12.88**	**<.001**	**1.26**	**1.72**
Performance Validity				**0.42**	**0.12**	**0.17**	**3.50**	**<.001**	**0.18**	**0.65**
Symptom Validity * Performance Validity				−0.10	0.22	−0.02	−0.44	.662	−0.54	0.34

Note. Cognitive performance = aggregate factor combining all cognitive performance measures utilized in the present study, derived from the exploratory factor analysis; symptom report = aggregate factor combining symptom report measures utilized in the present study, derived from the exploratory factor analysis; symptom validity = pass-fail status on any of the symptom validity measures at their respective cutoffs (MBIAS > 8, SIMS > 23, Validity-10 > 16); performance validity = pass-fail status on any of the performance validity measure (MSVT or b Test); R^2 = coefficient of determination; B = unstandardized beta; *SEB* = standard error of beta; β = standardized beta; LLCI = lower limit confidence interval; ULCI = upper limit confidence interval;

*
indicates interaction effect; bold font indicates statistical significance after adjustment for familywise error using False Discovery Rate (FDR).