


BRIEF COMMUNICATION

Patterns of Cognitive Test Scores and Symptom Complaints in Persons with TBI Who Failed Performance Validity Testing

Mark Sherer^{1,*} , Angelle M. Sander^{1,2}, Jennie Ponsford³, Leia Vos^{4,5}, Julia M.P. Poritz⁶, Esther Ngan⁷ and Luis Leon Novelo⁷

¹TIRR Memorial Hermann, Houston, TX, USA

²H. Ben Taub Department of Physical Medicine and Rehabilitation, Baylor College of Medicine & Harris Health System, Houston, TX, USA

³Monash-Epworth Rehabilitation Research Centre, Epworth Healthcare, Turner Institute of Brain and Mental Health, School of Psychological Sciences, Monash University, Melbourne, VIC, Australia

⁴Acute Mental Health, Neuropsychology, Zablocki (Milwaukee) VA Medical Center, Milwaukee, WI, USA

⁵Department of Psychiatry and Behavioral Medicine, Medical College of Wisconsin, Milwaukee, WI, USA

⁶Department of Obstetrics and Gynecology, University of Texas Medical Branch, Galveston, TX, USA

⁷Department of Biostatistics and Data Science, School of Public Health, University of Texas Health Science Center, Houston, TX, USA

(RECEIVED September 30, 2019; FINAL REVISION February 26, 2020; ACCEPTED March 6, 2020; FIRST PUBLISHED ONLINE May 27, 2020)

Abstract

Objective: To determine clinically meaningful subgroups of persons with traumatic brain injury (TBI) who have failed performance validity testing. **Method:** Study participants were selected from a cohort of 674 participants with definitive medical evidence of TBI. Participants were those who failed performance validity testing (the Word Memory Test, using the standard cutoffs). Participants were administered cognitive tests and self-report questionnaires. Test and questionnaire results were summarized as 12 dimension scores. Cluster analysis using the k-means method was performed. **Results:** Cluster analysis for the 143 retained participants indicated three subgroups. These subgroups differed on patterns of scores. Subgroup 1 was impaired for memory and had no excessive complaints. Subgroup 2 had impaired memory and processing speed as well as concern regarding cognition function. Subgroup 3 showed impairment on all cognitive tests and excess complaints in multiple areas. **Conclusions:** These results provide a preliminary basis for improved understanding of poor performance validity.

Keywords: Performance validity, Traumatic brain injury, Cognitive tests, Self-report questionnaires, Subgroups

Inclusion of performance validity tests (PVTs) as part of neuropsychological evaluations is now the standard of care in clinical neuropsychology (Bush et al., 2005; Heilbronner et al., 2009; Larrabee, 2012). These measures supplement standard cognitive tests by providing an indication of the extent to which the respondent gave a full effort to obtain valid scores. Failed performance validity testing generally indicates that the respondent did not give a full effort on cognitive testing. Note that failed performance validity is not synonymous with malingering as the latter requires the “presence of a substantial external incentive” (Slick, Sherman, & Iverson, 1999). While PVTs are most relevant to objective cognitive testing, persons who fail PVTs may report excessive complaints on self-report measures (Lippa, Pastorek, Romesser, et al., 2014).

Researchers have identified factors that may influence PVT performance. Such factors include malingering, somatoform disorder, psychological distress, perceived unfairness of the evaluation, boredom, and inadequate language proficiency among others (Bashem, Rapport, Miller, et al., 2014; Greher & Wodushek, 2017; Lippa, 2018). Henry et al. (2018) found that patient health beliefs were associated with validity test failure. Patients who endorsed beliefs that cognitive effort makes symptoms worse were more likely to show poor performance validity. Patients who endorsed beliefs that symptoms are due to an illness or injury as opposed to being normal experiences were also more likely to show poor performance validity. Thus, failure on PVTs is complex and different factors may result in a range of patterns of results (Chafetz, Williams, Ben-Porath, et al., 2015).

In a treatment context, the clinician has ongoing responsibility for the patient even if the patient shows poor performance validity. Some persons with clear medical

*Correspondence and reprint requests to: Mark Sherer, TIRR Research Center, 1333 Moursund, Houston, TX 77030. E-mail: Mark.Sherer@memorialhermann.org

documentation of moderate or severe traumatic brain injury (TBI) fail PVTs. Some of these individuals have true neurocognitive deficits (Larrabee, 2014) that may contribute to poor performance validity. Dismissal of the need for treatment in such cases can be deleterious; yet, it can be challenging to develop an appropriate treatment plan. In many cases, persons without true cognitive impairment may be experiencing emotional distress related to an injury. Of course, in some cases, persons referred for treatment may simply be malingering for secondary gain. Thus, the interpretation of failure on PVT tests can be complex, and the interaction of cognitive impairment, emotional distress, and motivational factors must be considered.

Investigations of patterns of cognitive test results indicate that persons who show invalid performance have cognitive test results that are distinct from persons who give valid effort (Frazier, Youngstrom, Naugle, et al., 2007). Also, as found by Lippa et al. (2014), persons with invalid performance on cognitive tests report more complaints on self-report measures. These findings suggest that there may be different patterns of cognitive test scores and self-report among persons with invalid performance.

In our previous work (Sherer, Sander, Nick, et al., 2015; Sherer, Nick, Sander, et al., 2017; Sherer, Ponsford, Hicks, et al., 2017), we identified 12 dimensions of signs and symptoms of TBI that can be derived from the 18 tests and questionnaires. The combinations of test scores and questionnaire results that composed each dimension are described in the Method section. Cluster analysis of the 12 dimension scores for 504 persons with TBI identified 5 subgroups (Sherer, Nick, et al., 2017). Scores on the 12 dimensions were plotted to generate profiles of TBI subgroups. These five subgroups differed on cognitive performance, subjective complaints, environmental supports, and performance validity. Almost all persons with failed performance validity were clustered in the same subgroup so that it was not possible to examine patterns of dimension scores within subgroups with invalid performance.

Morin and Axelrod (2017) examined subclasses (subgroups) of a large cohort of participants who were administered both cognitive tests and measures of emotional functioning and personality. Four subgroups were identified, with members of one subgroup generally showing invalid performance, markedly impaired cognitive performance, and numerous emotional complaints. The three other subgroups had low rates of poor performance validity with less extreme scores on cognitive tests and fewer emotional complaints. As with the Sherer et al. (2017) study, this analysis did not reveal subgroups of persons with invalid performance.

The present investigation sought to extend prior findings by identifying subgroups of persons who showed poor performance validity. We carried out this aim by studying a cohort of persons with TBI who failed performance validity testing when administered a battery of cognitive tests, self-report measures, and scales assessing environmental supports. We hypothesized that clinically meaningful subgroups with distinct patterns of test and questionnaire performance could be identified. We expected that the clinical profiles of these subgroups would have implications for clinical management.

METHOD

Participants

This study is a secondary analysis of a dataset originally developed for a study that identified dimensions of objective cognitive test scores and self-reported cognitive, emotional, and physical symptoms as well as environmental supports (Sherer, et al., 2015; Sherer, Nick, et al., 2017). Participants in the original studies were recruited at rehabilitation centers in Houston, Birmingham, and Detroit in the USA as well as, Melbourne, Australia (Sherer, Ponsford, et al., 2017). This research was carried out in accord with all relevant human subjects' protection guidelines including the Helsinki Declaration. The research was reviewed and approved by ethics committees at all participating sites including approval by the Baylor College of Medicine IRB at the primary study site.

Participants who were included for study: (a) had definitive, medical documentation of TBI occurring greater than 6 months prior to assessment, (b) were 18 to 64 years old, (c) had capacity to give informed consent, and (d) had ability to complete study measures in English. Diagnosis of TBI was based on history of head trauma with one or more of (a) observed loss of consciousness, (b) post-traumatic amnesia (PTA), and (c) trauma-related findings on CT scan. Persons with a preexisting medical or psychiatric condition that would affect performance on the assessment were excluded from study. For this investigation, only those persons who failed performance validity testing were retained.

Procedures

For the parent study for this secondary analysis, medical and research records were examined for demographic (age, education, and race/ethnicity) and injury characteristics (PTA, length of stay, first available Glasgow Coma Scale (GCS)). Participants completed 36 cognitive tests and self-report questions. Using data reduction techniques and cluster analyses, 12 dimensions of participant experience after TBI were derived from 18 measures. These dimensions and the tests/questionnaires that defined them were as follows: (1) Memory – Wechsler Adult Intelligence Scale – IV Letter-Number Sequencing (Wechsler, 2008), Rey Auditory Verbal Learning Test (RAVLT) total words for learning trials 1–5 (Rey, 1958), (2) Cognitive Processing Speed – Trail Making Test Part A (Reitan & Wolfson, 1985), Wechsler Adult Intelligence Scale – IV Coding (Wechsler, 2008), (3) Verbal Fluency – FAS (Gladsjo et al., 1999), (4) Self-reported Cognitive Symptoms – Traumatic Brain Injury-Quality of Life (TBI-QOL; Tulskey et al., 2016) Cognition-General Concerns, (5) Independence and Self-esteem – TBI-QOL Self-evaluation, TBI-QOL Independence, (6) Resilience – TBI-QOL Resilience, (7) Emotional Distress – TBI-QOL Anxiety, TBI-QOL Emotional and Behavioral Dyscontrol, (8) Post-concussive Symptoms – Neurobehavioral Symptom Inventory (Cicerone & Kalmar, 1995), (9) Physical Symptoms – TBI-QOL Headache, TBI-QOL Pain Interference, (10) Physical Functioning – TBI-QOL Upper Extremity,

(11) Economic and Family Support – Economic Quality of Life (Tulsky et al., 2015), Family Assessment Device General Functioning (Epstein, Baldwin, & Bishop, 1983), and (12) Performance Validity – Word Memory Test (WMT; Green, 2007). A more detailed description of data extracted from records, cognitive tests, questionnaires, and the PVT can be obtained by referring to Sherer, Sander, Nick, et al. (2015).

For neuropsychological tests, raw scores were converted to standardized scores utilizing adjustment for age, education, sex, and race/ethnicity as available in the WAIS-IV (Wechsler, 2008), Heaton, Miller, Taylor, and Grant (2004), and Schmidt (1996) norm sets. For the TBI-QOL, look-up tables were used to convert raw scores to T-scores (Tulsky et al., 2016). The reference sample for TBI-QOL measures was composed of persons with TBI. A similar look-up table was used for the Economic Quality of Life scale. Raw scores were used for the Neurobehavioral Symptom Inventory and the Family Assessment Device. An index of performance validity was calculated by averaging the scores from the three easy subtests of the WMT as recommended by Green (2007). Persons retained for this investigation were those who “failed” the WMT based on standard cutoffs.

Statistical Analysis

Given the similarity in results with regard to demographic and injury characteristics as well as subgroup profiles shown in an earlier investigation (Sherer, Ponsford, et al., 2017), the US and Australian samples were combined. Dimension scores from the combined sample were submitted to K-means cluster analysis with Euclidean distance. The number of clusters was selected using both variance ratio criterion (Caliński & Harabasz, 1974) and “global max” gap statistics (Tibshirani, Walther, and Hastie, 2001). Both approaches indicated that three was the optimal number of clusters. Profiles for subgroups were created by plotting the 12 mean dimensions scores for persons in each subgroup. Note that dimension scores were based on normalized z-scores for the 18 tests and questionnaires for the original cohort of 504. Summary statistics for identified subgroups were calculated for demographic and injury severity data as well as for the individual test and questionnaire scores for the 18 study measures.

RESULTS

Study Sample Descriptions and Comparisons

The total cohort included 170 persons with TBI from Australia and 504 from the USA for a combined sample of 674. There were 21 persons with missing WMT scores (USA $n = 13$, Australia $n = 8$) resulting in a sample size of 653. Of these, 143 (USA $n = 117$, Australian $n = 26$) showed poor performance validity as indicated by failure of the WMT as determined by application of the standard cutoffs. These 143 participants formed the study cohort for this investigation.

K-means cluster analysis of the 12 dimension scores for the 143 participants resulted in identification of 3 clusters (subgroups). Subgroups were labeled 1, 2, and 3 for

convenience. Note that for each dimension score, lower scores indicated poorer performance (e.g., more impaired cognitive abilities, more complaints on self-report measures). Comparisons for demographic, injury severity, and performance validity values for the three subgroups are shown in Table 1. The three subgroups did not differ on gender, race, injury severity (GCS, duration of PTA), age, years of education, or time since injury. While all three subgroups showed invalid performance, a difference was found in the overall WMT score ($p < .001$). Tukey’s multiple comparison testing showed that subgroup 3 showed worse overall performance on the WMT than subgroups 1 or 2.

Profile plots for the three subgroups are presented in Figure 1. Persons in subgroup 1 and subgroup 2 scored similarly on cognitive tests. However, subgroup 2 participants reported higher levels of complaints on cognitive, emotional, physical, and environmental support dimensions. Subgroup 3 participants scored more poorly on cognitive tests and reported more complaints than either subgroup 1 or subgroup 2.

To further demonstrate between group differences, we examined mean scores for the individual cognitive tests and self-report measures. Using greater than one standard deviation below the normative mean as indicating impairment, examination of the scores for the RAVLT Trials 1–5 revealed that all three subgroups were markedly impaired with scores greater than two standard deviations below the normative mean. Subgroup 1 scored grossly within normal limits for all remaining cognitive tests. Subgroup 2 also showed impairment on the Wechsler Adult Intelligence Scale – IV Coding subtest. Subgroup 3 obtained scores that were impaired or markedly impaired on all cognitive tests. For self-report measures, subgroup 1 scored similarly to a large cohort of persons with TBI from the TBI-QOL calibration sample (Tulsky et al., 2016). Subgroup 2 scored similarly to other persons with TBI on most self-report measures, but reported concern about cognition and elevated neurobehavioral complaints. Subgroup 3 showed high levels of concern regarding cognitive function, anxiety, and emotional and behavioral control. This group also reported disruption of activities by pain and decreased upper extremity function. They reported a moderate level of neurobehavioral complaints. These results are shown in Table 1 including Anova’s comparing the three groups, pairwise comparisons, and key effect sizes.

DISCUSSION

The primary aim of this investigation was achieved as distinct subgroups of persons with failed performance validity were identified. These three subgroups differed in clinically meaningful ways on cognitive performance and symptom complaints. One group scored within one standard deviation of the normative mean on all cognitive tests except for the test that is clearly a memory test (RAVLT Trials 1–5 Total) and within one standard deviation of the mean scores from TBI-QOL calibration sample on all TBI-QOL measures.

Table 1. Comparison of the three subgroups on demographics, injury severity, cognitive tests, and questionnaires

	Subgroup 1 (n = 43)	Subgroup 2 (n = 55)	Subgroup 3 (n = 45)	Anova		
	N (%)			p-value		
Categorical descriptors						
Gender	N (%)					
Male	37 (86%)	42 (76%)	33 (73%)	.317		
Female	6 (14%)	13 (24%)	12 (27%)			
Race	N (%)					
White	28 (65%)	26 (47%)	25 (56%)	.211		
Non-White	15 (35%)	29 (53%)	20 (44%)			
GCS group	N (%)					
Mild (13–15)	12 (28%)	17 (31%)	13 (29%)	.666		
Moderate (9–12)	4 (9%)	7 (13%)	5 (11%)			
Severe (3–8)	16 (37%)	17 (31%)	15 (33%)			
Intubated	10 (23%)	6 (11%)	3 (7%)			
Medicated	1 (2%)	3 (6%)	3 (7%)			
Missing	0 (0%)	5 (9%)	6 (13%)			
Continuous descriptors						
	Mean (SD)	Mean (SD)	Mean (SD)	Anova		
				p-value		
Age (y)	41.5 (13.7)	40.2 (10.9)	40.4 (12.9)	.863		
Years of education (y)	12.6 (2.5)	12.2 (2.2)	11.7 (2.1)	.233		
Time since injury (y)	5.9 (6.2)	7.4 (6.9)	7.0 (7.9)	.559		
PTA duration (d)	39.5 (39.8)	31.1 (37.6)	33.4 (34.3)	.579		
Measures						
	Mean (SD)	Mean (SD)	Mean (SD)	Anova p-value	Tukey comparisons	Cohen's <i>d</i> 1 vs. 3, 2 vs. 3
Letter-Number Sequencing	7.7 (3.1)	7.7 (2.7)	5.5 (2.8)	.001	1, 2 > 3	.7; .8
RAVLT Trials 1–5	26.5 (12.7)	29.3 (16.2)	23.5 (13.0)	.132	N/A	N/A; N/A
Trails A	43.3 (10.8)	41.9 (10.3)	36.7 (13.6)	.019	1 > 3	.5; N/A
Coding	7.2 (2.7)	6.6 (2.1)	5.4 (2.6)	.005	1 > 3	.6; N/A
Verbal Fluency (FAS)	41.7 (11.4)	45.0 (9.9)	33.2 (11.9)	.001	1, 2 > 3	.7; 1.1
TBI-QOL Cognition	43.5 (6.8)	34.1 (4.6)	28.4 (4.9)	.001	1 > 2 > 3	2.6; 1.2
TBI-QOL Self-esteem	55.6 (6.6)	47.0 (6.3)	39.7 (7.8)	.001	1 > 2 > 3	2.2; 1.0
TBI-QOL Independence	56.9 (7.5)	45.1 (5.2)	41.5 (6.9)	.001	1 > 2 > 3	2.1; .6
TBI-QOL Resilience	55.7 (8.8)	46.2 (5.7)	41.5 (8.0)	.001	1 > 2 > 3	1.7; .7
TBI-QOL Anxiety	47.8 (8.3)	58.1 (7.6)	64.9 (6.8)	.001	1 < 2 < 3	–2.3; –.9
TBI-QOL Emot Behav Dys	47.7 (7.4)	55.2 (8.1)	63.3 (10.1)	.001	1 < 2 < 3	–1.8; –.9
Neurobehavioral Symptom	.5 (3)	1.2 (4)	2.0 (6)	.001	1 < 2 < 3	–3.3; –1.6
TBI-QOL Headache	43.6 (7.7)	53.4 (7.5)	57.5 (7.4)	.001	1 < 2 < 3	–1.8; –.5
TBI-QOL Pain Interference	47.6 (7.9)	59.6 (6.4)	63.7 (10.2)	.001	1 < 2 < 3	–1.7; –.5
TBI-QOL Upper Extremity	49.0 (7.9)	43.2 (6.4)	38.2 (6.2)	.001	1 > 2 > 3	1.5; .8
Economic Quality of Life	60.4 (11.6)	50.4 (10.2)	47.2 (10.9)	.001	1 > 2, 3	1.2; N/A
FAD general functioning	1.8 (.4)	2.1 (.5)	2.4 (.5)	.001	1 < 2 < 3	–1.2; –.5
Word Memory Test	78.9 (9.7)	78.1 (8.7)	67.6 (13.3)	.001	1, 2 > 3	1.0; .9

Abbreviations: GCS = Glasgow Coma Scale; SD = standard deviation; y = years; PTA = post-traumatic amnesia; d = days; RAVLT = Rey Auditory Verbal Learning Test; TBI-QOL = Traumatic Drain Injury Quality of Life; Emot Behav Dys = emotional and behavioral disturbance. For all cognitive tests, higher scores indicate more intact function. For the TBI-QOL Anxiety, Emotional and Behavioral Dyscontrol, Headache, and Pain Interference scales, higher scores indicate greater severity of symptoms. Also, for the Neurobehavioral Symptom Inventory and Family Assessment Device, higher scores indicate greater severity of symptoms. For all other symptom measures, higher scores indicate more intact function. Pairwise comparisons were only reported when the overall Anova was significant and effect sizes were only reported when the pairwise comparison was significant. As subgroup 3 obtained the most impaired scores for all cognitive measures and the most symptomatic scores for all self-report measures, this group was used as the reference group for calculation of the effects sizes (Cohen's *d*). Effects sizes $\leq .20$ are small, $> .20$ but less than $.80$ moderate, and $\geq .80$ large (Cohen, 1988).

A second group scored in the impaired range for memory and the Wechsler Coding subtest while reporting much greater concern about cognitive function than the TBI-QOL calibration sample. A final group scored in the impaired range on all cognitive tests and reported substantial concern regarding

cognitive function, self-esteem, and upper extremity motor function on the TBI-QOL.

While the cluster profiles provide information regarding various patterns of cognitive scores and complaints, they do not reveal factors that might contribute to these different

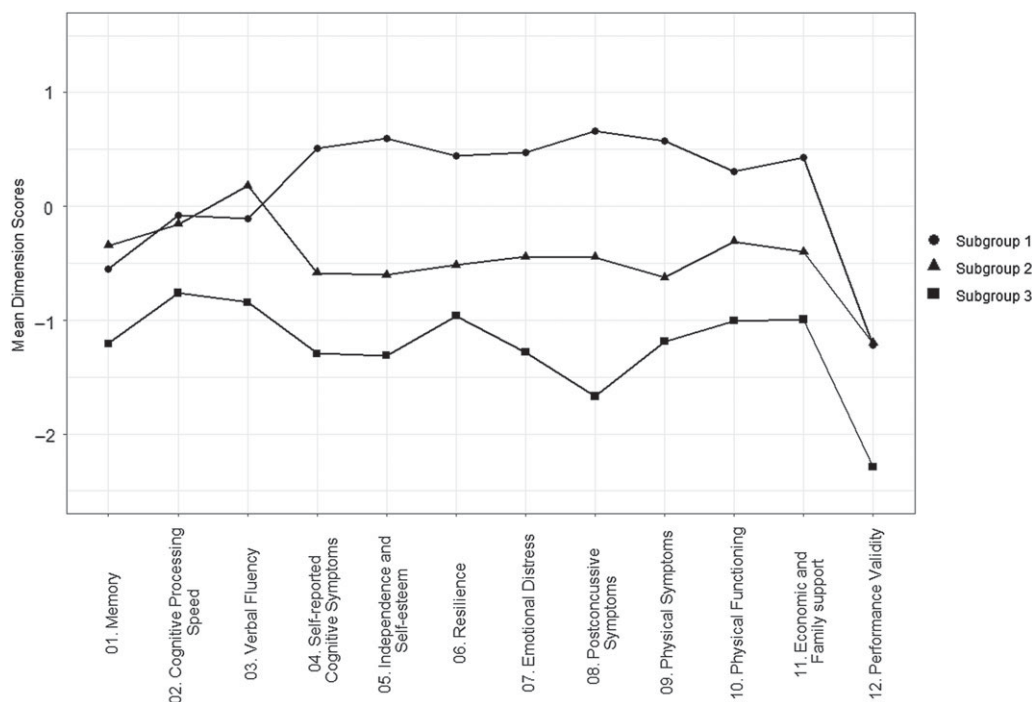


Fig. 1. Profile plots for the three subgroups.

profiles. Clinical examination of persons with failed performance validity should include a comprehensive history of factors, such as litigation status, applications for disability compensation, pre-injury psychiatric history, family, or other relationship dynamics that might provide secondary gain, apparent change in personal adjustment, and/or degree of psychological distress following injury. Test-taking behavior should be carefully observed to detect evidence of boredom or irritation with testing procedures.

Clinical intervention with persons who fail performance validity may be challenging. It is possible that factors that contribute to invalid performance may make patients more difficult to engage in treatment. Clinicians should be careful to listen to and acknowledge the patient's point of view. Presentation of medical information that shows that patient complaints are improbable should be carried out with attention to the patient's ability and willingness to entertain this information.

While admittedly speculative, we offer possible treatment implications of the subgroup profiles. For subgroup 1, interventions might focus on providing the sorts of recommendations that are applicable for the general population. These could include memory and organizational strategies using smart phone or other technologies as well as a focus on regular exercise, proper diet, and sleep hygiene depending on the patient's complaints. The emphasis should be on self-management of perceived impairments and complaints.

For subgroup 2, the initial focus of intervention could be on emotional distress using education or psychotherapy. Treatments that help persons with TBI to cope with their emotional distress and work toward realistic, value-oriented goals may be helpful.

Subgroup 3 members are likely to be the most difficult to engage in treatment. Here, creation of a strong therapeutic alliance with family or close others may be helpful. The social influence that can be brought to bear by the patient's social support network may be effective in enhancing engagement in treatment.

The treatment recommendations above are based on the clinical experience of the authors and require validation by additional investigation. Before this effort, the subgroup structure identified in this investigation should be cross-validated by additional research.

The present investigation has several limitations. The study cohort was drawn from persons with TBI who were living in the community and were not seeking treatment. The average interval from injury to study evaluation was 6.8 years. The majority of study participants sustained moderate or severe TBI rather than mild TBI. Information regarding whether participants were engaged in litigation or disability claims related to their injuries was not obtained. Persons with deficits due to prior neurologic illness or injury and those with severe psychiatric disturbance were excluded from this study. These factors may limit the applicability of study findings to other groups of persons with TBI and failed performance validity. Participants were not administered symptom validity tests and these tests may have contributed to understanding of the patterns of scores on subjective complaints. In addition, for this investigation, participants were only administered one performance validity measure. Best practice for clinical evaluations is administration of multiple measures of performance validity to provide converging evidence regarding adequacy of effort and to avoid reliance on a single measure

that might be failed due to actual cognitive impairment rather than poor performance validity (Board of Directors, 2007).

The findings of this investigation are an initial step toward a more sophisticated and individualized approach to managing patients with medically proven TBI and invalid performance on neuropsychological evaluation. Effective management of these persons may lead to improved quality of life for the patient and family/close others as well as decreased burden on healthcare resources. Further investigation is needed to build on this initial step. Future studies would benefit from examining participants who were seeking treatment.

ACKNOWLEDGMENTS

This research was partially supported by a grant from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR grant number 90DPTB0016). NIDILRR is a Center within the Administration for Community Living (ACL), Department of Health and Human Services (HHS). The contents of this article do not necessarily represent the policy of NIDILRR, ACL, HHS, and you should not assume endorsement by the Federal Government.

CONFLICT OF INTEREST

All authors declared no conflict of interest regarding this manuscript.

REFERENCES

- Bashem, J.R., Rapport, L.J., Miller, J.B., Hanks, R.A., Axelrod, B.N., & Millis, S.R. (2014). Comparison of five performance validity indices in bona fide and simulated traumatic brain injury. *The Clinical Neuropsychologist*, 28(5), 851–875.
- Board of Directors. (2007). American Academy of Clinical Neuropsychology (AACN) practice guidelines for neuropsychological assessment and consultation. *The Clinical Neuropsychologist*, 21, 209–231.
- Bush, S.S., Ruff, R.M., Troster, A.I., Barth, J.T., Koffler, S.P., Pliskin, N.H., . . . Silver, C.H. (2005). Symptom validity assessment: Practice issues and medical necessity. *Archives of Clinical Neuropsychology*, 20(4), 419–426.
- Caliński, T. & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics*, 3(1), 1–27.
- Chafetz, M.D., Williams, M.A., Ben-Porath, Y.S., Bianchini, K.J., Boone, K.B., Kirkwood, M.W., Larrabee, G.J., & Ord, J.S. (2015). Official position of the American Academy of Clinical Neuropsychology Social Security Administration Policy on validity testing: Guidance and recommendations for change. *The Clinical Neuropsychologist*, 29(6), 723–740.
- Cicerone, K.D. & Kalmar, K. (1995). Persistent postconcussion syndrome: The structure of subjective complaints after mild traumatic brain injury. *Journal of Head Trauma Rehabilitation*, 10(3), 1–17.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Epstein, N.B., Baldwin, L., & Bishop, D.S. (1983). The McMaster Family Assessment Device. *Journal of Marital and Family Therapy*, 9, 171–180.
- Frazier, T.W., Youngstrom, E.A., Naugle, R.I., Haggerty, K.A., Busch, R.M. (2007). The latent structure of cognitive symptom exaggeration on the Victoria Symptom Validity Test. *Archives of Clinical Neuropsychology*, 22, 197–211.
- Gladsjo, J.A., Schuman, C.C., Evans, J.D., Peavy, G.M., Miller, S.W., & Heaton, R.K. (1999). Norms for letter and category fluency: Demographic corrections for age, education, and ethnicity. *Assessment*, 6, 147–178.
- Green, P. (2007). The pervasive influence of effort on neuropsychological tests. *Physical Medicine and Rehabilitation Clinics of North America*, 18, 43–68.
- Greher, M.R. & Wodushek, T.R. (2017). Performance validity testing in neuropsychology: Scientific basis and clinical application – a brief review. *Journal of Psychiatric Practice*, 23(2), 134–140.
- Heaton, R.K., Miller, S.W., Taylor, M.J., & Grant, I. (2004). *Revised Comprehensive Norms for an Expanded Halstead-Reitan Battery: Demographically Adjusted Neuropsychological Norms for African American and Caucasian Adults*. Lutz, FL: Psychological Assessment Resources.
- Heilbronner, R.L., Sweet, J.J., Morgan, J.E., Larrabee, G.J., Millis, S.R., & Conference Participants. (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.
- Henry, G.K., Heilbronner, R.L., Suhr, J., Gornbein, J., Wagner, E., & Drane, D.L. (2018). Illness perceptions predict cognitive performance validity. *Journal of the International Neuropsychological Society*, 24(7), 735–745.
- Larrabee, G.J. (2012). Performance validity and symptom validity in neuropsychological assessment. *Journal of the International Neuropsychological Society*, 18(4), 625–630.
- Larrabee, G.J. (2014). Assessment of malingering. In G.J. Larrabee (Ed.), *Forensic Neuropsychology: A Scientific Approach* (2nd ed.) (pp. 116–159). New York: Oxford.
- Lippa, S.M. (2018). Performance validity testing in neuropsychology: a clinical guide, critical review, and update on a rapidly evolving literature. *The Clinical Neuropsychologist*, 32(3), 391–421.
- Lippa, S.M., Pastorek, N.J., Romesser, J., Linck, J., Sim, A.H., Wisdom, N.M., & Miller, B.I. (2014). Ecological validity of performance validity testing. *Archives of Clinical Neuropsychology*, 29(3), 236–244.
- Morin, R.T. & Axelrod, B.N. (2017). Use of latent class analysis to define groups based on validity, cognition, and emotional functioning. *The Clinical Neuropsychologist*, 31(6–7), 1087–1099.
- Reitan, R.M. & Wolfson, D. (1985). *The Halstead-Reitan Neuropsychological Test Battery: Theory and Clinical Interpretation*. Tucson, Arizona: Neuropsychology Press.
- Rey, A. (1958). *L'examen Clinique en psychologie*. Paris, France: Presses Universitaires de France.
- Schmidt, M. (1996). *Rey Auditory Verbal Learning Test (RAVLT): A Handbook*. Los Angeles: Western Psychological Services.
- Sherer, M., Nick, T.G., Sander, A.M., Melguizo, M., Hanks, R., Novack, T.A. . . . Tang, X. (2017). Groupings of persons with traumatic brain injury: A new approach to classifying traumatic brain injury in the post-acute period. *Journal of Head Trauma Rehabilitation*, 32(2), 125–133.

- Sherer, M., Ponsford, J., Hicks, A., Leon-Novelo, L., Ngan, E., & Sander, A.M. (2017). Cross-validation of a classification system for persons with traumatic brain injury in the posthospital period. *Journal of Head Trauma Rehabilitation, 32*(5), E17–E25.
- Sherer, M., Sander, A.M., Nick, T. G., Melguizo, M.S., Tulskey, D.S., Kinsala, P., Hanks, R., & Novack, T.A. (2015). Key dimensions of impairment, self-report, and environmental supports in persons with traumatic brain injury. *Rehabilitation Psychology, 60*(2), 138–146.
- Slick, D.J., Sherman, E.M.S., & Iverson, G.L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist, 13*(4), 545–561.
- Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society, Series B, 63*(2), 411–423.
- Tulskey, D.S., Kinsala, P.A., Lai, J.-S., Carlozi, N., Hammel, J., & Heinemann, A.W. (2015). Developing an item bank to measure economic quality of life for individuals with disabilities. *Archives of Physical Medicine and Rehabilitation, 96*(4), 604–613.
- Tulskey, D.S., Kinsala, P.A., Victorson, D., Carlozzi, N., Bushnik, T., Sherer, M., . . . Cella, D. (2016). TBI-QOL: Development and calibration of items banks to measure patient reported outcomes following traumatic brain injury. *Journal of Head Trauma Rehabilitation, 31*(1), 40–51.
- Wechsler, D. (2008). *Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV)*. San Antonio, Texas: Pearson.