Assessment
2014, Vol. 21(5) 562–569

DOI: 10.1177/1073191114520626
asm.sagepub.com
SAGE

The Relationship Between the Self-Report BASC-2 Validity Indicators and Performance Validity Test Failure After Pediatric Mild Traumatic Brain Injury

John W. Kirk[1], Christa F. Hutaff-Lee[2], Amy K. Connery[1], David A. Baker[1], and Michael W. Kirkwood[1]

Abstract

In adult populations, research on methodologies to identify noncredible performance and exaggerated symptoms during neuropsychological evaluations has grown exponentially in the past two decades. Far less work has focused on methods appropriate for children. Although several recent studies have used stand-alone performance validity tests with younger populations, a near absence of pediatric work has investigated other indices to identify response bias. The present study examined the relationship between the validity scales from the self-report Behavior Assessment System for Children, Second Edition (BASC-2) and performance on the Medical Symptom Validity Test (MSVT), a stand-alone performance validity test. The sample consisted of 274 clinically referred patients with mild traumatic brain injuries aged 8 through 17 years. Fifty patients failed the MSVT based on actuarial criteria. The majority of these patients (92%) provided valid self-report BASC-2 profiles, with only three patients (6%) producing an invalid profile due to an elevated F index. Analysis of valid/invalid self-report BASC-2 profiles and MSVT pass/fail did not reveal a significant relationship (p = 0.471, two-tailed Fisher's exact test). These findings suggest that performance validity tests like the MSVT provide substantively different information about the validity of a neuropsychological profile than that provided by the self-report validity scales of the BASC-2.

Keywords

BASC-2 Validity Scales, MSVT, symptom validity testing, performance validity testing, response bias, postconcussion, pediatric mild traumatic brain injury

In the adult neuropsychological scientific literature, research regarding methodologies to identify noncredible performance and exaggerated or feigned symptomatology has grown significantly in the past two decades (Boone, 2007; Larrabee, 2007; Sweet, King, Malina, Bergman, & Simmons, 2002). Both the National Academy of Neuropsychology and American Academy of Clinical Neuropsychology have now published position papers regarding the importance of evaluating performance and symptom validity, response bias, and malingering in the context of both independent and clinical neuropsychological contexts (Bush et al., 2005; Heilbronner, Sweet, Morgan, Larrabee, & Millis, 2009).

Larrabee (2012) recently suggested using the term *symptom validity* to refer to the accuracy of complaints on self-report measures and the term *performance validity* to refer to the validity of ability during task performance, assessed either by stand-alone tests or through atypical performance on standard neuropsychological tests (e.g., embedded measures). Among adult neuropsychologists, the use of both symptom validity tests (SVTs) and performance validity tests (PVTs) has become accepted practice across evaluative contexts (Sharland & Gfeller, 2007). Over the past 5 to 10 years, interest has also grown rapidly in the use of PVTs in pediatric populations.

A number of studies have now clearly demonstrated that certain stand-alone PVTs can be used appropriately with pediatric populations (Kirkwood, 2012). For example,

[1]Department of Physical Medicine & Rehabilitation, Children's Hospital Colorado and University of Colorado School of Medicine, Aurora, CO, USA
[2]Department of Pediatrics, Section of Child Neurology, Children's Hospital Colorado and University of Colorado School of Medicine, Aurora, CO, USA

Corresponding Author:
John W. Kirk, Children's Hospital Colorado, Concussion Program, B285, 13123 E. 16th Ave., Aurora, CO 80045, USA.
Email: drjohnkirk@comcast.net

pediatric patients as young as 5 or 6 years old can pass the commonly used Test of Memory Malingering (TOMM; Constantinou & McCaffrey, 2003; Donders, 2005; Kirk et al., 2011; MacAllister, Nakhutina, Bender, Karantzoulis, & Carlson, 2009; Tombaugh, 1996). The Victoria Symptom Validity Test (VSVT) measures recognition and retention of number sequences, although a recent study indicated that it can also be used as an effort measure in children as young as 6 years old (Brooks, 2012). The Word Memory Test (WMT) and Medical Symptom Validity Test (MSVT) require familiarity and skill with word recognition and are therefore inappropriate for very young children but have been demonstrated to be valid in children who have a second- to third-grade reading level (Blaskewitz, Merten, & Kathmann, 2008; Carone, 2008; Green & Flaro, 2003; Kirkwood & Kirk, 2010).

Multiple studies have found that a small but consistent percentage of general pediatric clinical patients fail PVTs, with failure rates ranging from 3% to 10% (Brooks, Sherman, & Krol, 2012; Carone, 2008; Donders, 2005; Kirk et al., 2011; MacAllister et al., 2009). Two other studies suggest that under certain conditions, rates of failure on PVTs in children are likely to be considerably higher. Kirkwood and Kirk (2010) found that 17% of clinically referred pediatric patients with mild TBI failed a stand-alone PVT. Chafetz, Abrahams, and Kohlmaier (2007) found an even higher percentage of children (28% to 37%) who failed a PVT in a compensation-seeking context (i.e., Social Security Disability benefit evaluations).

Kirkwood, Yeates, Randolph, and Kirk (2012) provided data to indicate that PVT performance in children has significant implications for interpreting the rest of the neuropsychological battery, not unlike that seen in adult populations. In this study, performance on the MSVT accounted for nearly 40% of the variance across neuropsychological tests in a sample of pediatric patients with mild TBI. Participants who failed the MSVT also performed significantly worse on nearly all neuropsychological tests including those measuring nonverbal reasoning, memory, attention, processing speed, and fine motor functioning. Effect sizes were large across most standardized tests, comparable to those seen in similar studies of adults, including samples with financial incentive to perform poorly (Constantinou, Bauer, Ashendorf, Fisher, & McCaffrey, 2005; Lange, Iverson, Brooks, & Rennison, 2010).

In contrast to this burgeoning pediatric research focused on stand-alone PVTs, a near absence of work has focused on the use of SVTs in pediatric populations. Symptom validity indictors embedded in self-report questionnaires measure the accuracy of endorsed symptomatic complaints and are time efficient and resistant to coaching. A wealth of literature has investigated negative response bias on adult self-report personality questionnaires. The bulk of this literature has focused on the Minnesota Multiphasic Personality Inventory, Second Edition (MMPI-2), and the MMPI-2-RF (restructured form), which contain validity indices that have been found to be quite effective in detecting feigned or exaggerated somatic, psychiatric, and cognitive complaints.

The relationship between SVTs and PVTs has been well-studied in adult clinical populations. In fact, the Response Bias Scale (RBS) of the MMPI-2 was developed to assist in detecting cognitive response bias associated with PVT failure on the WMT, as well as exaggerated memory or other cognitive complaints (Gervais, Ben-Porath, Wygant, & Green, 2007). In a retrospective analysis of neuropsychological assessments at a VA Hospital, the RBS and the Henry-Heilbroner Index of the MMPI-2 were found to be strong predictors of TOMM failure (Whitney, Davis, Shepard, & Herman, 2008). Furthermore, the Lees-Haley Fake Bad Scale (FBS) of the MMPI-2 and Symptom Validity–Revised of the MMPI-2-RF have been found to have significant correlations with passing or failing of stand-alone PVTs (Greiffenstein, 2010; Sellbom, Wygant, & Bagby, 2012; Youngjohn, Wershba, Stevenson, Sturgeon, & Thomas, 2011). Performance validity test failure on WMT, MSVT, and/or TOMM has also been found to be strongly associated with significant elevations on all MMPI-2 RF validity scales and several substantive scales such as Cognitive Complaints (Gervais, Wygant, Sellbom, & Ben-Porath, 2011; Jones, Ingram, & Ben-Porath, 2012).

In child and adolescent samples, substantially less attention has been given to the general validity of self-report profiles and the relationship between invalid self-report and noncredible performance as measured by PVTs. Indeed, among pediatric populations, only a few studies have evaluated self-report response bias, and none of these studies has investigated the relationship between self-report validity scales and PVT performance. Furthermore, the studies that have been completed have focused on feigned psychopathology rather than exaggerated health or cognitive symptoms, the latter of which are more likely to be reported during neuropsychological rather than psychiatric evaluations. To date, pediatric research on this topic has mostly involved adolescent simulator studies that have focused on the utility of the F (Infrequency) Scale, the L (Lie) Scale, and the F-K (Infrequency minus Defensiveness) scale from the Minnesota Multiphasic Personality Inventory–Adolescent (MMPI-A) to detect feigned psychopathology (Lucio, Duran, Graham, & Ben-Porath, 2002; Rogers, Hinds, & Sewell, 1996; Stein, Graham, & Williams, 1995). One study also investigated the clinical utility of the Validity (VAL) scale (e.g., highly improbable items), Inconsistency (INC) scale (e.g., comparing responses from pairs of highly correlated statements), Dissimulation (FB) scale (e.g., statements that were infrequent and characteristic of intentional distortion), and the Defensiveness (DEF) scale (e.g., overly positive self-presentation) from the

Personality Inventory for Youth (PIY; Wrobel et al., 1999). This study investigated the PIY profiles of adolescent patients admitted to a psychiatric hospital asked to minimize psychiatric problems or high school psychology students asked to exaggerate psychiatric symptoms. Significant elevations on the DEF scale were noted for the majority of patients asked to minimize psychological concerns, as well as elevated FB and VAL scales when students were asked to feign psychopathology. When high school students were asked to create a random response style, significant elevations were noted on the INC scale. The authors also found that the FB and DEF scales were significantly correlated with negative and positive self-description adjectives, respectively. The results provided evidence for using these validity scales to exaggerated, inconsistent, or defensive response styles on the self-report PIY form.

The Behavior Assessment Scale for Children–Second Edition (BASC-2) is a standardized behavioral rating form with teacher, parent, and self-report measures for children and adolescents (Reynolds & Kamphaus, 2004). Similar to the MMPI-A and the PIY, the self-report BASC-2 provides validity scales designed to assist with interpretation of responses. Of these, the F index includes items that are rarely endorsed by typically developing children and may also be considered a "fake bad" index as children may endorse items in an effort to look severely disturbed (Reynolds & Kamphaus, 2004). No identified studies have investigated the utility of the F index or any of the other validity indicators from the self-report BASC-2.

In the present study, we set out to determine if there was a relationship between the BASC-2 validity indicators and PVT performance in a sample of real-world pediatric patients. Based on previous literature, primarily from adult populations, we hypothesized that children who failed a PVT would be more likely to have elevations on BASC-2 validity indices, most notably the F index, suggesting a negative response bias.

Method

Participants

The sample consisted of 274 school-aged children and adolescents referred consecutively for outpatient clinical neuropsychological consultation following a mild TBI. Participants were drawn from a 4-year series of consecutive clinical cases referred to an outpatient concussion program at a children's hospital in the Rocky Mountain region of the United States. Patients were considered eligible for participation if they were administered the self-report format of the BASC-2, were aged 8 through 17 years at the time of evaluation, were within 1 year of sustaining a blunt head trauma, and were referred because of concerns or questions about the effects of underlying brain injury. A subgroup of this case series overlapped with participants presented in previously published studies (Kirkwood, Hargrave, & Kirk, 2011; Kirkwood & Kirk, 2010; Kirkwood et al., 2012). Participants from previous studies were included in the current sample if they had completed BASC-2 self-report questionnaires. The most common causes of injury in the current sample were recreation or sports-related (58%), falls (17%), motor vehicle–related trauma (10%), recreational vehicle (e.g., bicycle, skateboard; 6%), automobile versus pedestrian collision (3%), physical assault (2%), and other (4%). Children who had intracranial pathology on neuroimaging were included if their Glasgow Coma Scale score was never less than 13. Exclusionary criteria were forensic referral, neurosurgical intervention, injury resulting from nonaccidental trauma, and nontraumatic brain injury such as hypoxia, cerebrovascular insult, or infectious illness. If a patient was evaluated more than one time, only the first encounter data were used. The final sample included 274 participants. The sample was 60% male. Background and injury characteristics of the sample are provided in Table 1. Demographic and testing information was obtained through record review with all identifying information removed.

Measures

The BASC-2 is a measure of social, emotional, and behavioral functioning, which has parent, teacher, and self-report formats. The self-report format has five validity scales to help the clinician determine the quality of the response profile. The F Index is a tally of the number of items in which the patient gave an overly negative self-report (i.e., "faking bad"). The F Index is focused primarily on items that might be useful in detecting feigned psychopathology, rather than exaggerated health or cognitive complaints. The Consistency Index identifies cases in which the patient gave differing responses to items that are usually answered similarly (i.e., measure of random responding). The Response Pattern Index is the tally of the number of times an item response differs from the response to a similar previous item (i.e., measure of attention to item content). The L Index is the tally of the number of times the patient responded *True* or *Almost Always* to an unrealistically positive self-description or *False* or *Never* to a mildly self-critical statement that most people endorse (i.e., "faking good"). The V Index is the tally of the nonsensical items that are rarely endorsed if the patient is paying close attention to and understands the item content. Each scale generates categorical scales of "acceptable," "caution," or "extreme caution." The BASC-2 self-report has two different versions; one is for children 8 to 11 years old (Child Form), which includes 139 items, and the other is for those 12 to 21 years old (Adolescent Form), which includes 176 items. Both versions include all five validity indicators, which are calculated and reported in a

Table 1. Background and Injury Characteristics of Participants.

Total participants	*N* = 274
Age (years)	*M* = 14.72, *SD* = 2.28
Grade	*M* = 8.8, *SD* = 2.3
Male	*n* = 163 (60%)
Caucasian	*n* = 230 (84%)
Estimated full scale IQ[a]	*M* = 102.9, *SD* = 12.4
Maternal years of education	*M* = 15.0, *SD* = 2.2
Paternal years of education	*M* = 15.2, *SD* = 2.6
Premorbid history of attention-deficit/hyperactivity disorder	*n* = 49 (18%)
Premorbid history of diagnosed learning disability	*n* = 29 (11%)
Premorbid history of special education services	*n* = 39 (14%)
No premorbid history of ADHD, LD, or special education services	*n* = *198* (72%)
Weeks since injury	*M* = 10.5, *SD* = 9.8, *Mdn* = 7.0
Loss of consciousness	*n* = 52 (19%)
Neuroimaging conducted	*n* = 202 (73%)
Intracranial findings identified by CT or MRI	*n* = 24 (12%)
Families in or planning litigation	*n* = 23 (8%)
Families seeking disability compensation	*n* = 0 (0%)
Participants charged with a crime	*n* = 0 (0%)

a. Based on performance of the 264 participants administered the two subtest Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999).

similar format (Reynolds & Kamphaus, 2004). Most (89%) BASC-2 self-reports were from the Adolescent Form, which is not surprising given that the mean age of the sample was 14.7 years.

The MSVT is a computerized forced-choice verbal memory test designed to evaluate effort and memory. The primary effort indices are the Immediate Recognition (IR), Delayed Recognition (DR), and Consistency (CNS) scores. The test requires about 5 minutes of direct administration time (i.e., not including the delay time between IR and DR). A list of 10 semantically related word pairs is presented two times on a computer screen. Examinees are then asked to choose the correct word from pairs consisting of the target and a foil, during IR and DR conditions. After each response, examinees receive auditory and visual feedback. Examinees are then asked to recall the words during paired associate (PA) and free recall (FR) conditions. Participants in the current project were administered the MSVT in a standardized fashion, except that the examiner stayed in the room during the entire administration. The actuarial criteria proposed by Green (2004) were considered indicative of noncredible performance.

The Wechsler Abbreviated Scales of Intelligence is an individually administered assessment of estimated cognitive ability (Wechsler, 1999). For the present study, the two subtest version was used, which yields an estimated Full Scale IQ Index score.

Procedure

The project was reviewed and approved by a university-affiliated institutional review board. Patients underwent testing at the earliest 1 week postinjury and at the latest 52 weeks postinjury. The median testing time was 7 weeks postinjury. Most children underwent an abbreviated battery of neuropsychological tests rather than a more comprehensive evaluation (as discussed in Kirkwood et al., 2008), with the MSVT and BASC-2 included in the battery of tests. The battery also included the Wechsler Abbreviated Scales of Intelligence (two subtest version) for the majority (96%) of the sample. The MSVT was administered to all participants within the first hour of the evaluation.

Results

Cutoff scores for the MSVT were determined in accordance with the criteria outlined in the test manual (Green, 2004). Participants were separated into two groups based on their performance on the MSVT. Participants who achieved passing scores on all of the primary effort indices on the MSVT were placed within the "pass" group, while participants failing at least one or more of the three effort indices were placed in the "fail" group. Of the 274 participants, 50 (18%) failed at least one of the three primary effort indices. This percentage is consistent with the number of patients who were judged to have provided noncredible effort in an earlier subgroup of the same case series, after possible false positives and false negatives on the MSVT were taken into account (Kirkwood & Kirk, 2010).

Performance profiles of participants who failed the MSVT and those who did not are provided in Table 2. As would be expected, the MSVT pass group performed significantly better than the MSVT fail group across all 5 measured MSVT variables. The two groups did not differ in

Table 2. MSVT Performance in Those Passing and Failing the Primary Effort Indices.

	MSVT pass (n = 224)				MSVT fail (n = 50)			
	M	*SD*	*Mdn*	Range	*M*	*SD*	*Mdn*	Range
Age (years)	14.78	2.27	15.38	8.00-17.92	14.41	2.33	14.96	8.42-17.83
Grade	8.85	2.31	9.00	2-12	8.62	2.26	9.00	3-12
MSVT IR	99.51	1.70	100.00	90-100	76.90	18.07	80.00	25-100
MSVT DR	99.15	2.21	100.00	90-100	67.40	16.97	70.00	20-100
MSVT CNS	98.88	2.48	100.00	90-100	68.80	15.14	72.50	35-95
MSVT PA	98.57	4.31	100.00	70-100	63.40	24.40	70.00	0-100
MSVT FR	76.10	12.48	80.00	35-100	48.70	17.55	50.00	10-90

Note. MSVT = Medical Symptom Validity Test; IR = Immediate Recognition; DR = Delayed Recognition; CNS = Consistency scores; PA = Paired Associates; FR=Free Recall.

Table 3. Comparison of Self-Report BASC-2 Validity Scales in Those Passing and Failing MSVT.

	MSVT		
BASC-2 SRP validity scales	Pass (number of participants)	Fail (number of participants)	Significant test (two-tailed, Fisher's exact test)
F scale			
Invalid	4	3	p = .117
Within normal limits	220	47	
Response Pattern			
Invalid	1	1	p = .332
Within normal limits	223	49	
Consistency Scale			
Invalid	10	0	p = .217
Within normal limits	214	50	
L Scale			
Invalid	14	0	p = .081
Within normal limits	210	50	
V Scale			
Invalid	2	0	p = 1.00
Within normal limits	222	50	
Any validity scale			
Invalid	29	4	p = .471
Within normal limits	195	46	

Note. BASC-2 = Behavior Assessment System for Children, Second Edition; MSVT = medical symptom validity test.

gender, age, grade, estimated Full Scale IQ score, reading grade level, ethnic/racial status (classified as Caucasian or other), parent education, history of premorbid LD, ADHD, or reading problems, time since injury, or whether the injury was associated with loss of consciousness or neuroimaging pathology. Litigation status of the groups did not differ either. A total of 23 families reported that they were engaged in or planning litigation, with only 3 of these patients failing the MSVT. Thus, assuming honest reporting, a maximum of 3 out of the 50 patients in the MSVT fail group (6%) were potentially driven by compensation seeking, as compared to 20 out of the 224 in the MSVT pass group (9%).

The BASC-2 validity data were dichotomized, with valid results classified as 0 and either "caution" or "extreme caution" classified as 1. Some patients produced a response profile with more than one elevated validity scale. All profiles were considered "valid" if all five scales were valid, and "invalid" if any of the five validity scales were reported with "caution" or "extreme caution." Due to the number of cell counts with less than five observed cases, a Fisher's exact test was calculated on the BASC2 validity scales for MSVT pass/fail performance. These results did not yield any significant correlations between invalid responding on the BASC2 and MSVT classification (see Table 3).

Among the 50 patients who failed the MSVT, only 4 (8%) generated a profile that was deemed invalid, whereas 29 of the 224 (13%) who passed the MSVT had elevations on BASC-2 validity scales that rendered their profiles

invalid. Only 3 patients (6%) in the MSVT fail group produced an invalid profile due to an elevated F index, as compared to the MSVT pass group where 4 patients (2%) had an elevated F index.

Discussion

The current study supports previous research findings that adolescents and children as young as 8 years of age are capable of noncredible performance during neuropsychological evaluation, even in clinical contexts when secondary gain may not be readily apparent. In this relatively large homogenous case series, 18% of patients performed below the actuarial cutoffs on the MSVT. This was comparable to the number of patients who were judged to have provided noncredible effort in an earlier version of the same case series once possible false positives and false negatives were taken into account (Kirkwood & Kirk, 2010). Although children and adolescents are clearly capable of feigning cognitive symptoms in pursuit of financial gain, compensation-seeking behavior did not drive the majority of MSVT failures in this clinical sample. At the time of the neuropsychological contact, no cases were seeking disability compensation, and only 6% of the participants who failed the MSVT were engaged in or planning litigation, compared with 9% of those who passed the MSVT. The question of why children failed PVTs was not the focus of the current study. Suffice it to say, however, the clinicians evaluating the study participants judged the reasons to be quite varied and to include attempts in order to obtain external gains (e.g., additional support at school) and fulfill internal psychological needs (e.g., somatoform disorder). Certain children were also judged to be simply noncompliant. These and many other possible explanatory factors for noncredible effort in children have been discussed elsewhere in a separate case-based analysis (Kirkwood, Kirk, Blaha, & Wilson, 2010).

The primary purpose of the current study was to determine if there was a relationship between the BASC-2 validity indicators and a measure of performance validity in a sample of children and adolescents with mild TBI. Although the overlap between symptom validity and performance validity tests has been extensively researched in adult samples, few studies have investigated the use of SVTs in pediatric populations, with existing studies focused on feigned psychopathology rather than complaints of somatic or cognitive symptoms. No identified studies have investigated the relationship between validity indicators on a self-report questionnaire and performance on a stand-alone PVT in a pediatric population.

Based on previous research with adults, we hypothesized that children who failed a PVT would be more likely to have elevations on BASC-2 validity indices, most notably the F index. However, results of the current study did not find any relationship between invalidity as measured by the BASC-2 and that measured by the MSVT. In our sample, scoring in the invalid range on any of the BASC2 validity scales occurred infrequently, with only 8% of the sample producing an invalid protocol (i.e., any of the five validity scales indicating "caution" or "extreme caution"). Among the invalid protocols, only 3 patients (6%) in the MSVT fail group produced an invalid profile due to an elevated F index, as compared to the MSVT pass group where 4 patients (2%) had an elevated F index. Reynolds and Kamphaus (2004) note that F index items were typically endorsed by less than 3% of the respondents in the item-development normative samples. At least compared to the normative samples, those individuals in the MSVT fail group produced a slightly higher rate of elevated F index items; however, this elevation was not correlated with performance on the MSVT. Furthermore, there were no significant correlations among those individuals who passed/failed the Consistency Index, Response Pattern Index, L Index, or V index and those who passed/failed the MSVT.

These findings stand in stark contrast to published studies with adult samples, where invalid responding on SVTs relates strongly to PVT failure (Dionysus, Denney, & Halfaker, 2011; Iverson, Henrichs, Barton, & Allen, 2002; Whitney et al., 2008). Closer analysis of the item content on the BASC-2 reveals that the F index contains items primarily focused on extreme psychiatric complaints, with only one somatic item (e.g., dizziness) and no cognitively focused items, whereas the MMPI validity scales focus on somatic, psychiatric, and cognitive symptoms. This difference between the BASC-2 F index and FBS and RBS may help explain the lack of correlation between MSVT failure and elevated F index among our sample of children with mild TBI.

The results of this study indicate that relying exclusively on the BASC-2 validity scales as an indication of the validity of the overall cognitive/neuropsychological data will almost certainly significantly underestimate the percentage of patients providing invalid data during evaluation. Thus, the results reinforce the value of adding objective PVTs to the neuropsychological evaluation of school-age patients. Although the five validity scales from the BASC-2 self-report may be important in the identification of valid questionnaire response profiles, they are likely measuring a different type of response validity than is identified through stand-alone PVTs.

The results of the present study need to be interpreted in the context of several limitations. The participants were drawn from a sample of convenience composed of children and adolescents for whom persistent questions or concerns were apparent following a mild TBI. Because most youth can be expected to recover relatively quickly after such an injury, the participants are unlikely to be representative of the majority of patients with mild TBI. The current sample

was also skewed toward adolescent Caucasian patients from well-educated families. Another limitation was that the MSVT was the only PVT administered to all patients. Like any classification decision that relies on a single test, decisions about noncredible performance based solely on the MSVT will include some false positive and false negative errors (Kirkwood & Kirk, 2010). Thus, select cases in the MSVT pass group likely provided noncredible performance during other aspects of the test battery, and select cases in the MSVT fail group likely provided credible performance. We also used validity scales predetermined by the test publisher and did not perform an item analysis on self-report BASC-2 questionnaire data. If we had, we may have been able to identify items that were more strongly associated with MSVT performance.

Despite these limitations, the work remains worthwhile, as it is the first identified study that has examined the relationship between self-report and performance validity in a clinical pediatric sample. Future studies need to examine the behavioral/emotional symptom clusters (e.g., BASC-2 self-report clinical index scales) in children who fail objective PVTs. Patients providing invalid performance-based data may present with a certain cluster of personality traits or emotional difficulties (e.g., somatizing style, elevated depressive symptoms) that could differentiate them from their peers who do not fail PVTs. Future research also needs to examine the value of the validity indices of other self-report measures (e.g., MMPI-A, Achenbach Youth Self-Report) that are commonly used during pediatric neuropsychological evaluations.

Authors' Note

A preliminary version of this article was presented at the annual meeting of the American Academy of Clinical Neuropsychology in Chicago, June 2010.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

Blaskewitz, N., Merten, T., & Kathmann, N. (2008). Performance of children on symptom validity tests: TOMM, MSVT, and FIT. *Archives of Clinical Neuropsychology*, *23*, 379-391.

Boone, K. B. (Ed.). (2007). *Assessment of feigned cognitive impairment: A neuropsychological perspective*. New York, NY: Guilford Press.

Brooks, B. L. (2012). Victoria symptom validity test performance in children and adolescents with neurological disorders. *Archives of Clinical Neuropsychology*, *27*, 858-868.

Brooks, B. L., Sherman, E. M., & Krol, A. L. (2012). Utility of TOMM Trial 1 as indicator of effort in children and adolescents. *Archives of Clinical Neuropsychology*, *27*, 23-29.

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 NAN policy & planning committee. *Archives of Clinical Neuropsychology*, *20*, 419-426.

Carone, D. A. (2008). Children with moderate/severe brain damage/dysfunction outperform adults with mild-to-no brain damage on the Medical Symptom Validity Test. *Brain Injury*, *22*, 960-971.

Chafetz, M. D., Abrahams, J. P., & Kohlmaier, J. (2007). Malingering on the Social Security disability consultative exam: A new rating scale. *Archives of Clinical Neuropsychology*, *22*, 1-14.

Constantinou, M., Bauer, L., Ashendorf, L., Fisher, J. M., & McCaffrey, R. J. (2005). Is poor performance on recognition memory effort measures indicative of generalized poor performance on neuropsychological tests? *Archives of Clinical Neuropsychology*, *20*, 191-198.

Constantinou, M., & McCaffrey, R. J. (2003). Using the TOMM for evaluating children's effort to perform optimally on neuropsychological measures. *Child Neuropsychology*, *9*, 81-90.

Dionysus, K. E., Denney, R. L., & Halfaker, D. A. (2011). Detecting negative response bias with the Fake Bad Scale, Response Bias Scale, and Henry-Heilbronner Index of the Minnesota Multiphasic Personality Inventory-2. *Archives of Clinical Neuropsychology*, *26*, 81-88.

Donders, J. (2005). Performance on the test of memory malingering in a mixed pediatric sample. *Child Neuropsychology*, *11*, 221-227.

Gervais, R., Ben-Porath, Y., Wygant, D., & Green, P. (2007). Development and validation of a Response Bias Scale (RBS) for the MMPI-2. *Assessment*, *14*, 196-208.

Gervais, R., Wygant, D., Sellbom, M., & Ben-Porath, Y. (2011). Association between symptom validity test failure and scores on the MMPI-2-RF validity and substantive scales. *Journal of Personality Assessment*, *93*, 508-517.

Green, P. (2004). *Manual for the Medical Symptom Validity Test*. Edmonton, Canada: Green's Publishing.

Green, P., & Flaro, L. (2003). Word memory test performance in children. *Child Neuropsychology*, *9*, 189-207.

Greiffenstein, M. (2010). The MMPI-2 symptom validity scale (FBS) not influenced by medical impairment: A large sleep center investigation. *Assessment*, *17*, 269-277.

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., & Millis, S. R. (2009). American Academy of Clinical Neuropsychology Consensus Conference Statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist, 23*, 1093-1129.

Iverson, G. L., Henrichs, T. F., Barton, E. A., & Allen, S. (2002). Specificity of the MMPI-2 Fake Bad Scale as a marker for personal injury malingering. *Psychological Reports*, *90*, 131-136.

Jones, A., Ingram, M., & Ben-Porath, Y. (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*, 790-815.

Kirk, J. W., Harris, B., Hutaff-Lee, C. F., Koelemay, S. W., Dinkins, J. P., & Kirkwood, M. W. (2011). Performance on the Test of Memory Malingering (TOMM) among a large clinic-referred pediatric sample. *Child Neuropsychology*, *17*, 242-254.

Kirkwood, M. W. (2012). Overview of tests and techniques to detect negative response bias in children. In E. Sherman & B. Brooks (Eds.), *Pediatric forensic neuropsychology* (pp. 136-161). New York, NY: Oxford University Press.

Kirkwood, M. W., Hargrave, D. D., & Kirk, J. W. (2011). The value of WISC-IV digit span subtest in detecting noncredible performance during pediatric neuropsychological evaluations. *Archives of Clinical Neuropsychology*, *26*, 377-384.

Kirkwood, M. W., & Kirk, J. W. (2010). The base rate of suboptimal effort in a pediatric mild TBI sample: Performance on the Medical Symptom Validity Test. *The Clinical Neuropsychologist*, *24*, 860-872.

Kirkwood, M. W., Kirk, J. W., Blaha, R. Z., & Wilson, P. (2010). Noncredible effort during pediatric neuropsychological exam: A case series and literature review. *Child Neuropsychology*, *16*, 604-618.

Kirkwood, M. W., Yeates, K. O., Randolph, C., & Kirk, J. W. (2012). The implications of symptom validity test failure for ability-based test performance in a pediatric sample. *Psychological Assessment*, *24*, 36-45.

Kirkwood, M. W., Yeates, K. O., Taylor, H. G., Randolph, C., McCrea, M., & Anderson, V. A. (2008). Management of pediatric mild traumatic brain injury: A neuropsychological review from injury through recovery. *The Clinical Neuropsychologist*, *22*, 769-800.

Lange, R. T., Iverson, G. L., Brooks, B. L., & Rennison, V. L. A. (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*, 961-972.

Larrabee, G. J. (Ed.). (2007). *Assessment of malingered neuropsychological deficits*. New York, NY: Oxford University Press.

Larrabee, G. J. (2012). Performance validity and symptom validity in neuropsychological assessment. *Journal of International Neuropsychological Society*, *18*, 625-631.

Lucio, E., Duran, C., Graham, J. R., & Ben-Porath, Y. S. (2002). Identifying faking bad on the MMPI-A with Mexican adolescents. *Assessment*, *9*, 62-69.

MacAllister, W. S., Nakhutina, L., Bender, H. A., Karantzoulis, S., & Carlson, C. (2009). Assessing effort during neuropsychological evaluation with the TOMM in children and adolescents with epilepsy. *Child Neuropsychology*, *15*, 521-531.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment System for Children–Second Edition manual*. Circle Pines, MN: AGS Publishing.

Rogers, R., Hinds, J. D., & Sewell, K. W. (1996). Feigning psychopathology among adolescent offenders: Validation of the SIRS, MMPI-A and SIMS. *Journal of Personality Assessment*, *67*, 244-257.

Sellbom, M., Wygant, D., & Bagby, M. (2012). Utility of the MMPI-2-RF in detecting non-credible somatic complaints. *Psychiatry Research*, *197*, 295-301.

Sharland, M. J., & Gfeller, J. D. (2007). A survey of neuropsychologists' beliefs and practices with respect to the assessment of effort. *Archives of Clinical Neuropsychology*, *22*, 213-223.

Stein, L. A. R., Graham, J. R., & Williams, C. R. (1995). Detecting fake-bad MMPI-A profiles. *Journal of Personality Assessment*, *65*, 415-427.

Sweet, J. J., King, J. H., Malina, A. C., Bergman, M. A., & Simmons, A. (2002). Documenting the prominence of forensic neuropsychology at national meetings and in relevant professional journals from 1990 to 2000. *Clinical Neuropsychology*, *16*, 481-494.

Tombaugh, T. N. (1996). *Test of memory malingering*. Toronto, Ontario, Canada: Multi-Health Systems.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: Harcourt Assessment.

Whitney, K., Davis, J., Shepard, P., & Herman, S. (2008). Utility of the Response Bias Scale (RBS) and other MMPI-2 validity scales in predicting TOMM performance. *Archives of Clinical Neuropsychology*, *23*, 777-786.

Wrobel, T. A., Lachar, D., Wrobel, N. H., Morgan, S. T., Gruber, C. P., & Neher, J. A. (1999). Performance of the Personality Inventory for Youth validity scales. *Assessment*, *6*, 367-376.

Youngjohn, J., Wershba, R., Stevenson, M., Sturgeon, J., & Thomas, M. L. (2011). Independent validation of the MMPI-2-RF somatic/cognitive and validity scales in TBI litigants tested for effort. *The Clinical Neuropsychologist*, *25*, 463-476.