Advances in Pediatric Neuropsychology Test Interpretation: Importance of Considering Normal Variability and Performance Validity

Brian Brooks, PhD &

Michael Kirkwood, PhD, ABPP/CN

Children's Hospital Colorado

Affiliated with University of Colorado Anschutz Medical Campus

Children's Hospital Colorado

Acknowledgements

Children's Hospital Colorado Team Collaborators
- David Baker, PsyD
- Amy Connery, PsyD, ABPP
- Robin Peterson, PhD, ABPP

- Tess Simpson, PsyD
- Mike Dichiaro, MD
- Amy Kanallakan, MD
- Jackie Murray, RN
- Wendy Pierce, MD
- Aaron Provance, MD
- Anne Stratton, MD
- Pam Wilson, MD

Student Collaborators
- Robert Blaha, MA
- Cassie Green, MA
- David Hargrave, PsyD
- Christa Hutaff-Lee, PhD
- Jennifer Larson, PhD

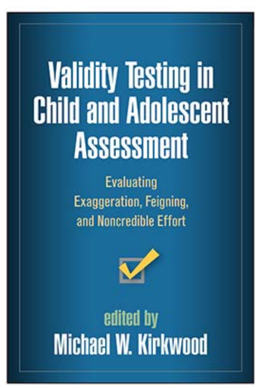
External Collaborators/Influences
- Brian Brooks, PhD
- Mike Chafetz, PhD, ABPP
- Jacobus Donders, PhD, ABPP
- John Kirk, PsyD, ABPP
- Bill MacAllister, PhD, ABPP/CN
- Chris Randolph, PhD, ABPP/CN
- Elisabeth Sherman, PhD
- Keith Yeates, PhD, ABPP/CN

Financial Disclosure

- No relevant conflicts of interest
  - Employment: Children's Hospital Colorado
  - Consulting: No financial payment for TBI roles. Have served on CDC Pediatric Mild TBI Expert Panel and the Mild TBI/Concussion Work Group, NINDS Common Data Elements Project. Have received payment for consulting work on several non-TBI related projects.
  - Consulting: Financial payment from PAR as expert reviewer during development of the Memory Validity Profile (MVP; Sherman & Brooks, 2016) and Children and Adolescent Memory Profile (ChAMP; Sherman & Brooks, 2015). No financial compensation from the sale of the tests.
  - Stock ownership: No relevant investments.
  - Research support: Principal investigator, co-investigator, or collaborator on pediatric TBI grants funded by numerous organizations including CDC, NIDRR, NIMH, Colorado TBI Trust Fund, Children's Hospital Colorado.
  - Honoraria: Have received honoraria and expense reimbursement from multiple institutions and professional scientific bodies for presenting pediatric TBI and validity testing work at meetings, conferences, and symposiums.
  - Royalties: I receive royalties on the sale of these books:
    - Kirkwood, M.W. & Yeates, K.O. (Eds.). (2012). Mild Traumatic Brain Injury in Children and Adolescents: From Basic Science to Clinical Management. New York: Guilford Press.
    - Stucky, K., Kirkwood, M.W., & Donders, J. (Eds.). (2014). Clinical Neuropsychology Study Guide and Board Review. New York: Oxford University Press.
    - Armstrong, K.S., Beebe, D.W., Hilsabeck, R.C., & Kirkwood, M.W. (2008). Board Certification in Clinical Neuropsychology: A Guide to Becoming ABPP/ABCN Certified without Sacrificing Your Sanity. New York: Oxford University Press.

Financial Disclosure

- Financial disclosure
  - Royalties: I receive royalties on a book that I edited on pediatric validity testing.



Guilford Press

I. Introduction

1. A Rationale for Performance Validity Testing in Child and Adolescent Assessment, Michael W. Kirkwood
2. Terminology and Diagnostic Concepts, Elisabeth M. S. Sherman
3. Understanding Deception from a Developmental Perspective, Eric Peterson & I
4. Performance and Symptom Validity: A Perspective from the Adult Literature, C

II. Detection Methods and Other Validity Test Usage Matters

5. Review of Pediatric Performance and Symptom Validity Tests, Michael W. Kirk
6. Clinical Strategies to Assess the Credibility of Presentations in Children, Domi
7. Motivations Behind Noncredible Presentations: Why Children Feign and How t David A. Baker & Michael W. Kirkwood
8. Managing Noncredible Performance in Pediatric Clinical Assessment, Amy K. Connery &Yana Suchy
9. Ethical Considerations in Pediatric Validity Testing, William S. MacAllister & Mars

III. Validity Testing across Evaluative Settings

10. Child and Adolescent Psychoeducational Evaluations, Allyson G. Harrison
11. Pediatric Clinical Neuropsychological Evaluations with Medical Populations, Bria
12. Pediatric Sports-Related Concussion Evaluations, Martin L. Rohling, Jennifer Lan Melissa M. Womble
13. Pediatric Forensic Neuropsychological Evaluations, Jacobus Donders
14. Disability: Social Security Supplemental Security Income Exams for Children, Michael D. Chafetz

In my very biased opinion, excellent group of authors and recent and relevant state of the science information so….

5

Objective Methods to Detect Noncredible Data

Performance Validity Tests (PVTs)
Used to detect inadequate effort or noncredible performance during testing

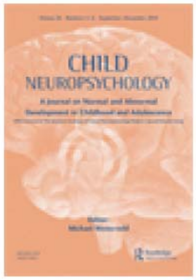
Symptom Validity Tests (SVTs)
Used to detect noncredible responding during self-report measures

aka, "effort" or "malingering" tests but movement away from these terms

Stand-Alone Tests

Embedded Indicators

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



A survey of neuropsychologists' use of validity tests with children and adolescents

Brian L. Brooks[1,2,3], Danielle M. Ploetz[1], and
Michael W. Kirkwood[4,5]

[1]Neurosciences Program (Brain Injury and Rehabilitation), Alberta Children's Hospital, Calgary, AB, Canada
[2]Departments of Paediatrics, Clinical Neurosciences, and Psychology, University of Calgary, Calgary, AB, Canada
[3]Behaviour and the Developing Brain Program, Alberta Children's Hospital Research Institute, Calgary, AB, Canada
[4]Children's Hospital Colorado, Aurora, CO, USA
[5]Department of Physical Medicine and Rehabilitation, University of Colorado School of Medicine, Aurora, CO, USA

A web-based survey of validity test use by North American neuropsychologists was conducted, with 282 participants meeting inclusion criteria. Respondents indicated that they use a median of one stand-alone performance validity test (PVT), one embedded PVT, and one symptom validity test (SVT) per pediatric assessment. The vast majority of respondents indicated they give at least one PVT (92%) and at least one SVT (88%) during each pediatric assessment. A meaningful difference in validity use (i.e., at least a medium effect size) was only found for those who engage in forensic work, with those clinicians giving more stand-alone PVTs than those who do not conduct forensic work. The most frequently used validity measures in pediatric assessments are presented, as are reasons participants reported for both using and not using validity tests. Limitations and qualitative comparisons to other surveys on validity test use with adults are discussed.

7

Table 2 Practice Characteristics of Respondents.

Characteristic	*n*	% of total sample that report any activity in this area
Number of Clinical Assessments Per Month	282 (mean = 8.5, *SD* = 7.0)	
Ages of Patients Seen for Assessments		
0–5 years	185	65.6
6–12 years	267	94.7
13–17 years	277	98.2
18+ years	233	82.6
Language for Assessments		
English (100% of the time)	238	84.4
Spanish (At least some of the time)	35	13.3
French (At least some of the time)	4	1.6
Other Languages (At least some of the time)	16	5.8
Professional Settings		
Private Practice	146	52.1
Hospital	177	63.0
Schools	10	3.6
Prison/Detention Centre	2	0.8
Psychiatric Facility	7	2.8
Academics	26	9.3
Other	12	4.3
Professional Activities		
Clinical Assessment	275	97.9
Forensic Medico-Legal	89	31.7
Therapy	72	25.6
Trainee Supervision	163	58.0
Research	127	45.2
Classroom Teaching	39	13.9
Administration	136	49.3
Other activity (e.g., didactics)	13	4.6

8

Table 3 Methods Used to Detect Invalid Data in those Under 18 years of Age.

Method (in Descending Order of Popularity)	Yes, I use this method (%)	No, I do not use this method (%)
Behavioral observations indicative of poor compliance	92.9	7.1
Discrepancies among records	90.8	9.2
Severity of cognitive impairment inconsistent with the condition	83.0	17.0
Pattern of cognitive impairment inconsistent with condition	81.9	18.1
Implausible self-reported symptoms in interview	79.4	20.6
Flagged validity scales in objective personality or behavioral measures	73.7	26.3
Score below empirical cutoffs on stand-alone measures of validity	73.4	26.6
Scores below chance on forced choice test	71.9	28.1
Implausible changes in test scores	65.8	34.2
Scores below empirical cutoffs on embedded measures	60.3	39.7
None	0.7	99.3

9

- Historically, reliance on subjective judgment to determine validity in pediatric evaluations
  - *"Mary appeared to put forth her best effort on all tasks. The results are therefore considered a reliable and valid representation of her cognitive functioning."*
- Objective instrumentation has allowed us to move away from subjective judgments in vast majority of other domains (e.g., attention, language, memory, mood). Why should test effort be different?
- Imagine with intelligence….
  - *"Mary appeared to have below average intelligence. The results therefore indicate that she has an intellectual disability (aka, mental retardation)."*

Problems with relying only on subjective judgment to identify noncredible data

- General literature suggests flaws in clinical judgment and decision-making
  - Ziskin & Faust (1988); Dawes (1994); Garb (1998)
- Two neuropsychologically-focused studies by Faust in 1988 (children and adolescents)
  - Youth (9-12; 15-17) told to perform less well than usual but not so obvious that the person testing them would know they were faking
    - No instruction in how to fake
  - Clinicians sent vignette that youth in MVC with LOC, unremarkable CT, and memory complaints some months later; clinicians asked to judge whether data abnormal and then speak to etiology
  - Majority of clinicians thought the profile reflected abnormality
  - Detection rate for malingering 0%
  - Majority of clinicians confident in their judgments
- Faust studies criticized (eg, clinicians have access to more than simply test results)
  - Bigler (1990); McCaffrey & Lynch (1992)
  - Yet, collectively, raise a number of questions
- Objective methodology has clear potential of reducing classification errors
  - In our experienced group in Denver, many cases would not be identified without PVTs

Consensus Need for Objective Methodology

Independent Evaluations

- NAN (2005)
  - *"Symptom exaggeration or fabrication occurs in a sizable minority of neuropsychological examinees, with greater prevalence in forensic contexts. Adequate assessment of response validity is essential in order to maximize confidence both in the results of ability measures and in the diagnoses and recommendations that are based on the results."*
- AACN (2009)
  - *"Especially because research has shown repeatedly that experienced experts are inaccurate in identifying valid versus invalid ability performances from mere observation of behavior or test scores, for a clinician to choose not to use effort tests and embedded validity indicators requires a solid justification, especially within a forensic context."*
- Sweet (2009)
  - *"In fact, failure to proactively assess for possible malingering in a forensic case is now considered below the standard of acceptable practice…"*

Clinical Evaluations

- NAN (2005)
  - *"Although the use of SVTs in clinical contexts may not always be indicated….determinations regarding the validity of patient performance are generally aided by the inclusion of SVTs in neuropsychological evaluations."*
- AACN (2009)
  - *"Even in a routine clinical context, the presence of problematic effort and response bias can potentially invalidate results. The assessment of effort and genuine reporting of symptoms is important in all evaluations."*

- Bill MacAllister and Marsha Vasserman (2015)
  - *The use of formal validity testing as part of the routine assessment of children and adolescents should no longer be considered optional, as it is in alignment with the professional guidelines of the field (e.g., NAN, AACN) and consistent with the ethical guidelines for psychologists (APA, 2002). Integration of performance validity data into neuropsychological practice reflects the current state of the field.*

MacAllister, W.S. & Vasserman, M. (2015). Ethical considerations in pediatric validity testing. In M.W. Kirkwood (Ed.), *Validity Testing in Child and Adolescent Assessment: Evaluating Exaggeration, Feigning, and Noncredible Effort*.

13

Objective Methods to Evaluate Validity

PVTs

- Stand-alone performance-based validity tests
  - Both forced and non-forced choice tests
  - Pros: designed specifically to maximize discriminability between groups so should have better classification statistics
  - Cons: battery time and money

- Indices from conventional tests ("embedded" indicators)
  - Simple cut-offs and atypical performance patterns
  - Pros: time and effort efficient, resistant to coaching, allow for more continuous monitoring of effort
  - Cons: classification statistics generally not as good as stand-alone tests

SVTs

- General behavioral/personality inventories

- Disorder-specific inventories

Table 5 Frequency Use of PVTs in Children and Adolescents.

PVT	Never (%)	Rarely (%)	Sometimes (%)	Often (%)	Almost Always (%)	
21-Item Test	93.8	3.3	2.9	0.0	0.0	
Amsterdam Short Memory Test	99.3	0.7	0.0	0.0	0.0	
Automatized Sequences Task	90.2	2.5	2.2	1.8	3.3	
The b Test	92.0	4.0	3.3	0.7	0.0	
CARB	95.3	2.2	1.8	0.4	0.4	
CVLT-C Discriminability Index	37.0	5.8	20.7	20.7	15.9	37%
CVLT-II Effort Algorithm Wolf 2010	86.6	4.0	5.8	1.8	1.8	
CVLT-II Forced Choice	27.2	7.2	29.0	21.4	15.2	37%
Dot Counting Test	87.0	3.6	6.5	2.9	0.0	
MSVT	62.0	6.9	14.1	10.9	6.2	17%
NV-MSVT	85.1	5.1	3.3	5.4	1.1	
Reliable Digit Span	34.8	8.0	13.8	22.1	21.4	44%
Rey-15 Item Test	66.3	17.0	9.1	6.2	1.4	
TOMM	22.1	12.0	31.2	20.7	14.1	35%
Word Completion Memory Test	95.7	2.2	0.7	1.4	0.0	
WMT	69.6	78.0	8.0	8.7	5.8	15%
VSVT	85.5	6.2	4.0	2.9	1.4	

15

Stand-Alone PVTs Investigated in Pediatric Populations

- Pediatric PVT Reviews
  - Kirkwood (2012). Overview of tests and techniques to detect negative response bias in children. In Sherman & Brooks (Eds.). (2012). *Pediatric Forensic Neuropsychology.*
  - DeRight & Carone (2013). Assessment of effort in children: A systematic review. *Child Neuropsychology.*
  - Kirkwood (2015). Review of PVTs and SVTs in children. In Kirkwood (Ed.). *Validity Testing in Child and Adolescent Assessment: Evaluating Exaggeration, Feigning, and Noncredible Effort.*

Table 2. Strength of empirical evidence estimates for the most commonly used stand-alone performance validity tests in pediatric populations

	Strength of Evidence for Use in Children			
	Community samples	Clinical samples	Secondary gain samples	Simulation samples
Computerized Assessment of Response Bias (CARB)	--	+	--	--
Dot Counting Test (DCT)	--	--	--	--
Fifteen Item Test (FIT)	+	+	--	--
Medical Symptom Validity Test (MSVT)	+	++	+	++
Memory Validity Profile (MVP)	++	+	--	+
Nonverbal Medical Symptom Validity Test (NV-MSVT)	--	+	--	--
Test of Memory Malingering (TOMM)	++	++	+	++
Victoria Symptom Validity Test (VSVT)	--	+	--	--
Word Memory Test (WMT)	+	++	--	+

Note. ++ adequate to strong evidence; + modest evidence; -- no or conflicting evidence

Adapted with permission. Kirkwood, M. W. (2012). Overview of tests and techniques to detect negative response bias in children. In E. M. S. Sherman & B. L. Brooks (Eds.), *Pediatric Forensic Neuropsychology* (pp. 136-161). New York: Oxford University Press.

Additional tests with potential utility but need more study

- Several PVTs have been investigated in only one identified pediatric study or by one group
- Of these…..

Most Promising

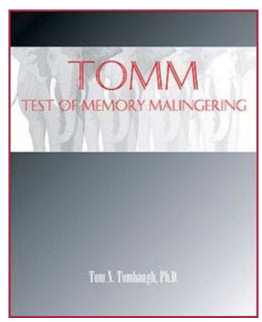
- **Nonverbal Medical Symptom Validity Test** (Green, 2008)
  - Green, Flaro, Brockhaus, & Montijo (2012); Harrison et al. (2014)
- **Victoria Symptom Validity Test** (Slick, Hopp, Strauss, & Thompson, 1997)
  - Brooks (2012)

Mixed Results or Very Little Work

- **Dot Counting Test** (Lezak, 1983; Rey, 1941)
  - Martin, Haut, Stainbrook, & Franzen (1995); Rambo et al. (2015)
- **21-Item Test** (Iverson, 1998)
  - Martin, Haut, Stainbrook, & Franzen (1995)
- **Computerized Assessment of Response Bias** (Allen, Conder, Green & Cox, 1997)
  - Courtney, Dinkins, Allen, & Kuroski (2003); Harrison et al. (2014)
- **Amsterdam Short-Term Memory Test** (Schmand & Lindeboom, 2004)
  - Rienstra, Spaan, & Schmand (2010)
- **Word Completion Memory Test** (WCMT; Hilsabeck & LeCompte, 1997)
  - Rienstra, Spaan, & Schmand (2010)

Test of Memory Malingering (TOMM)

- What is it?
  - Developed by Tombaugh (1996)
  - Examinee presented 50 line drawings twice; forced choice response during IR and DR, with optional retention trial



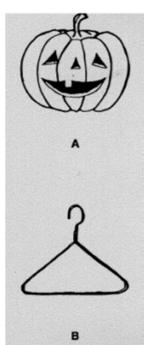


Table 3. Summary of pediatric studies focused on the Test of Memory Malingering

Source	Population	N	Age Range	Mean Age (SD)	Trial 1 Mean (SD)	Trial 2 Mean (SD)	% Passing*
Constantinou & McCaffrey (2003)	Cyprus Community	61	5 – 12	8.4 (2.1)	46.8 (3.4)	49.5 (1.7)	97%
Constantinou & McCaffrey (2003)	U.S. Community	67	5 – 12	7.9 (2.0)	45.9 (3.7)	49.9 (0.3)	100%
Rienstra et al. (2010)	Netherlands Community	48	7 – 12	9.9 (1.6)	--	50.0 (0.0)	100%
Schneider et al. (2014)	U.S. Community	30	4 – 7	5.6 (0.8)	43.3 (4.2)	47.1 (4.7)	85%**
Donders (2005)	U.S. Clinical mixed	100	6 – 16	11.9 (3.4)	46.5 (4.2)	49.7 (0.72)	97%
MacAllister et al. (2009)	U.S. Clinical epilepsy	60	6 – 17	~13.0 (~3.5)	43.5 (6.6)	47.5 (4.8)	90%
Kirk et al. (2012)	U.S. Clinical mixed	101	5 – 16	10.6 (3.2)	46.7 (3.2)	49.6 (0.9)	96%
Loughan & Perna (2012)	U.S. Clinical mixed	86	6 – 18	11.6 (3.2)	45.3 (5.6)	48.2 (4.0)	90%
Brooks et al. (2012)	U.S. Clinical mixed	53	6 – 19	12.4 (4.1)	44.0 (5.6)	48.4 (5.0)	94%
Ploetz et al. (in press)	U.S. Clinical mixed	266	5 – 18	13.0 (3.7)	46.9 (4.7)	46.9 (6.3)	94%
Schneider et al. (2014)	U.S. Clinical ADHD	36	4 – 7	5.5 (1.0)	41.1 (6.3)	44.4 (9.2)	85%**
Gast & Hart (2010)	U.S. Juvenile court	107	12 – 17	15.4 (1.4)	46.7 (3.4)	49.7 (0.9)	99%
Chafetz (2007)	U.S. Social Security Disability applicants	96	6 – 16	10.6 (2.7)	38.2 (5.5)	40.6 (2.4)	40%
Nagle et al. (2006)	U.S. Simulation controls	17	6 – 12	~8.6 (~2.9)	--	49.7 (0.8)	100%
Blaskewitz et al. (2008)	Germany Simulation controls	51	6 – 11	8.9 (1.0)	--	49.8 (0.9)	100%
Gunn et al. (2010)	Australia Simulation controls	50	6 – 11	~8.7 (~1.8)	46.6 (3.2)	49.2 (1.3)	98%
Rambo et al. (2013)	U.S. Simulation controls	17	6 - 12	10.1 (1.8)	45.7 (4.4)	49.8 (0.75)	100%

TOMM bottom line

- Most empirical work
- Likely appropriate with children 5+ years
- Appears specific in all but the most impaired children
- Relatively low cost
- Unlikely to be as sensitive as some other measures (Blaskewitz et al; Rambo et al; missed 1/3 simulators)
- More time consuming than some other PVTs

Kirkwood (2015). Review of PVTs and SVTs in children. In Kirkwood (Ed.). *Validity Testing in Child and Adolescent Assessment: Evaluating Exaggeration, Feigning, and Noncredible Effort*. Guilford Press.

Green's Word Memory Test (WMT) & Medical Symptom Validity Test (MSVT)

Boat Water

Dog Cat

- What are they?
  - WMT: Patient presented twice with 20 semantically linked words on computer
  - MSVT: Patient presented twice with 10 semantically linked words
  - Followed by a number of trials
  - Primary effort measures: IR, DR, and Consistency between two trials
  - Originally normed for adults but Flaro provided data from children with variety of clinical disorders
  - Profile analysis allows for examination of whether a fail is a "true impairment profile"

WMT

Table 4. Summary of pediatric studies focused on the Word Memory Test

Source	Population	N	Age Range	Mean Age (SD)	IR % Mean (SD)	DR % Mean (SD)	CNS %	% Passing*
Rienstra et al. (2010)	Netherlands Community	48	7 – 12	9.9 (1.6)	--	--	--	100%
Green et al. (2012)	Canada Clinical mixed ≥ 3rd grade reading level	380	--	13.4 (2.7)	95.9 (5.7)	95.9 (7.0)	93.8 (7.7)	90%
Courtney et al. (2003)	U.S. Clinical mixed – younger group	55	6 – 9	8.5 (1.2)	Average effort scores 74.2 (18.8)			--
Courtney et al. (2003)	U.S. Clinical mixed – older group	56	10 – 17	13.4 (2.0)	Average effort scores 93.4 (10.4)			--
Larochette & Harrison (2012)	U.S. Clinical Learning Disability	63	11 - 14	12.2 (0.6)				91%
Gunn et al. (2010)	Australia Simulation controls	50	6 – 11	~8.7 (~1.8)	90.6 (7.6)	95.3 (6.1)	--	98%

Kirkwood (2015). Review of PVTs and SVTs in children. In Kirkwood (Ed.). *Validity Testing in Child and Adolescent Assessment: Evaluating Exaggeration, Feigning, and Noncredible Effort*. Guilford Press.

MSVT

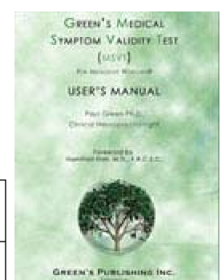


Table 5. Summary of pediatric studies focused on the Medical Symptom Validity Test

Source	Population	N	Age Range	Mean Age (SD)	IR % Mean (SD)	DR % Mean (SD)	CNS % Mean (SD)	% Passing *
Green et al. (2009)	Canada Community	56	7 – 11	9.2 (1.7)	98.6 (3.8)	98.6 (3.0)	97.6 (5.4)	96%
Green et al. (2009)	Brazil Community young	36	6 – 10	8.7 (1.4)	95 (5)	99 (3)	94 (8)	98%
Green et al. (2009)	Brazil Community old	34	11 – 15	12.4 (1.3)	96 (4)	100 (2)	96 (4)	
Green et al. (2012)	Canada Clinical mixed ≥ 3rd grade reading level	265	--	13.6 (2.9)	98.8 (3.7)	98.0 (4.3)	97.3 (5.8)	95%
Carone (2008)	U.S. Clinical mixed	38	--	11.8 (3.1)	98.6 (3.7)	97.6 (6.3)	96.7 (9.0)	95%
Kirkwood & Kirk (2010)	U.S. Clinical mild TBI	193	8 – 17	14.5 (2.4)	95.5 (5.3)	93.6 (5.4)	93.9 (4.8)	83%
Chafetz et al. (2007)	U.S. Social Security Disability applicants	25	6 – 16	11.5 (2.6)	86.4 (8.0)	84.2 (9.9)	87.8 (9.1)	37%
Blaskewitz et al. (2008)	Germany Simulation controls	51	6 – 11	8.9 (1.0)	98.6 (2.5)	99.6 (1.2)	98.2 (3.6)	98%

Kirkwood (2015). Review of PVTs and SVTs in children. In Kirkwood (Ed.). *Validity Testing in Child and Adolescent Assessment: Evaluating Exaggeration, Feigning, and Noncredible Effort*. Guilford Press.

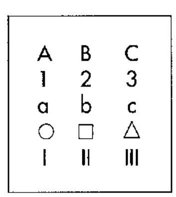
WMT & MSVT Bottom Line

- Good evidence appropriate with children with 3rd grade reading level or better
  - Solid specificity above this threshold
- Evidence to indicate more sensitive than TOMM (Blaskewitz et al.; Rambo et al.,)
  - Consistent with our experience in Denver
- Available multiple languages
- Potential added benefit of “profile analysis” to detect true impairment vs. noncredible effort
- MSVT can be administered quickly so good as screening measure
- MSVT cost per use; WMT annual fee (historically)

23

Rey Fifteen-Item Test (FIT)

- What is it?
  - Best known of Rey’s validity procedures
  - Patient shown 15 items and then asked to draw as many as can
  - Adapted by others – eg, Boone et al. (2002) developed a recognition format



Green, Kirk, Connery, Baker, & Kirkwood. Rey FIT after pediatric mild TBI (2014)

Trial	Cutoff score	Sensitivity (%)	Specificity (%)
FIT recall	<9	12	98
	<10	12	97
	<11	14	97
	<12	16	97
	<13	49	87
	<14	53	87
	<15	59	86
FIT recall with recognition trial	<22	25	97
	<23	29	96
	<24	39	95
	<25	47	92
	<26	55	91
	<27	59	89
	<28	63	83
	<29	67	80
	<30	71	66

TABLE 7.7. Mean Rey Fifteen Item Test (FIT) Scores, Standard Deviations, and Percentage Passing in Pediatric Studies

Source	Population	N	Age Range	Mean Age (SD)	Test Mean (SD)	% Passing
Constantinou & McCaffrey (2003)	Cyprus community	61	5–12	8.4 (2.1)	10.8 (4.7)	—
Constantinou & McCaffrey (2003)	U.S. community	67	5–12	7.9 (2.0)	10.8 (4.3)	—
Blaskewitz et al. (2008)	Germany simulation controls	51	6–11	8.9 (1.0)	12.6 (2.2)	100%

Bottom line

- Probably appropriate in higher functioning children 11+ years
- Extreme caution in younger/lower functioning children
- Blaskewitz et al. (2008) and Cassie Green et al. (2014) suggest traditional cutoff scores quite insensitive to noncredilbe effort
- Results from Green et al. (2014) support adding Boone recognition trial
  - increased sensitivity considerably, without altering specificity, at least among the higher functioning 8-17 year olds

Memory Validity Profile (Sherman & Brooks – 2016)

MVP Memory Validity Profile™

PAR

- First commercially available stand-alone PVT designed specifically for children/teens
  - Exciting development for all of us pediatric neuropsychologists
- Consists of verbal and visual paradigms
- Underwent test development like commercially produced cognitive tests including pilot testing, expert panel review, bias review, and refinement testing
- Normed on 1,200 US youth aged 5-21 years, 200 youth with clinical diagnoses, and 45 children in a simulation design study
- First validity test with age-adjusted cut scores to minimize false positives in young children
- Not yet available for independent review but certainly promising

26

Embedded Indicators

- Extensive literature in adult populations (Boone, 2007; Larrabee, 2007)
- Much less study in children

Embedded PVTs Investigated in Pediatric Populations

- Review
  - Kirkwood (2015). Review of PVTs and SVTs in children. In Kirkwood (Ed.). *Validity Testing in Child and Adolescent Assessment: Evaluating Exaggeration, Feigning, and Noncredible Effort*.

- **Digit Span**
  - Blaskewitz et al. (2008)
  - Kirkwood et al. (2011)
  - Welsh et al. (2012)
  - Loughan et al. (2012)
  - Perna et al. (2014)
  - Harrison & Armstrong (2014)
- **CVLT-C**
  - Baker et al. (2014)
  - Brooks et al. (2015)
- **Automatized Sequences Task**
  - Kirkwood et al. (2014)
- **Matrix Reasoning**
  - Kirkwood et al. (2012)
  - Rambo et al. (2013)
- **Symptom Validity Scale**
  - Chafetz et al. (2007; 2008)
- **CNS Vital Signs**
  - Brooks et al. (2014)
- **ChAMP**
  - Sherman & Brooks (2015)

Digit Span as an Embedded Indicator

- Adult studies
  - Dozens of studies across a variety of samples (Reviews: Suhr & Barrash, 2007; Babikian & Boone, 2007)
  - Age-corrected scaled scores
    - < 5 has typically been associated with > 90% specificity, with sensitivity ranging from about 25% to 50%
  - Reliable Digit Span (Greiffenstein, Baker, & Gola, 1994)
    - Calculated by summing the longest string of digits repeated without error over two trials under both forward and backward conditions
      - Eg, pass both trials 3 digits forward, pass both trials of two digits back = 5
    - Cutoff of < 8 or < 7 has produced sensitivity values above 50% in nearly all adult studies (specificity less ideal in more severely affected populations at this level)
    - Cutoff < 6, sensitivity is lowered to around 40-60% but specificity improves more consistently to at least 90%

- First child study: Blaskewitz, Merten, & Kathmann (2008)
  - German simulation design with 70 children (6 – 11 year olds)
  - WISC-III Digit Span subtest administered
  - Using adult cutoff for RDS, majority of matched controls (59%) failed
  - Classification statistics for lower RDS cutoff scores and other Digit Span scores not published

OXFORD
UNIVERSITY PRESS

Archives of Clinical Neuropsychology 26 (2011) 377–384

Archives of CLINICAL NEUROPSYCHOLOGY

The Value of the WISC-IV Digit Span Subtest in Detecting Noncredible Performance during Pediatric Neuropsychological Examinations†

Michael W. Kirkwood*, David D. Hargrave, John W. Kirk

Department of Physical Medicine & Rehabilitation, The Children's Hospital, Colorado and University of Colorado School of Medicine, Aurora, CO, USA

*Corresponding author at: The Children's Hospital, Concussion Program, B285, 13123 E. 16th Ave., Aurora, CO 80045, USA. Tel.: +720-777-6193; fax: +720-777-7297.
E-mail address: kirkwood.michael@tchden.org (M.W. Kirkwood).

Accepted 21 April 2011

Abstract

In adult populations, research on methodologies to identify negative response bias has grown exponentially in the last two decades. Far less work has focused on methods appropriate for children. Although several recent studies have demonstrated the appropriateness of using stand-alone symptom validity tests with younger populations, a near absence of pediatric work has investigated embedded validity indicators. The present study examined the classification value of several scores derived from the WISC-IV Digit Span subtest. The sample consisted of 274 clinically referred mild traumatic brain injury patients aged 8 through 16 years. Fourteen percent of the participants failed both the Medical Symptom Validity Test and Test of Memory Malingering, which was used as the criterion for noncredible effort. For age-corrected scaled scores, a score of ≤5 resulted in the optimal cut-score, yielding sensitivity of 51% and specificity of 96%. For Reliable Digit Span, the optimal cut-score was ≤6, with sensitivity of 51% and specificity of 92%. Although only moderately sensitive, Digit Span scores are likely to have good utility in identifying noncredible performance in relatively high-functioning older children and adolescents. Indeed, classification statistics produced in this pediatric sample compare favorably with those produced in many real-world adult patients.

Keywords: Digit span; Reliable digit span; Wechsler intelligence scale for children; Symptom validity testing; Response bias; Postconcussion; Mild traumatic brain injury

Table 3. Sensitivity and specificity values for various WISC-IV Digit Span score cutoffs

	Sensitivity %	Specificity %
Digit Span Scaled Score		
≤3	22	100
≤4	38	99
≤5	51	96
≤6	68	89
≤7	78	76
≤8	81	67
≤9	84	53
≤10	95	39
Reliable Digit Span		
≤4	14	100
≤5	27	100
≤6	51	92
≤7	76	69
≤8	89	49
≤9	97	24
≤10	100	15
Digit Span Forward Raw Score		
≤4	19	100
≤5	24	99
≤6	41	94
≤7	62	84
≤8	81	68
≤9	87	54
≤10	97	31
Digit Span Backward Raw Score		
≤3	8	100
≤4	24	97
≤5	51	91
≤6	76	70
≤7	84	48
≤8	95	28
≤9	100	18
≤10	100	10

31

- Denver results compare favorably to many real-world adult populations
  - Moderate sensitivity (~50%) when specificity > 90%
  - Loughan et al (2012) found similar classification statistics using a cut-score of ss ≤ 4 (Sens = 43%, Spec 91%)
    - Only 7 noncredible cases total though; 6/7 mild TBI cases
- Keep in mind nature of sample
  - Higher functioning older kids/teens with mild neurological injury
- Different results will almost certainly be obtained in lower functioning populations (e.g., those with neurologically or developmentally-based problems)
  - Indeed….

Digit Span as Embedded Indicator

OXFORD UNIVERSITY PRESS

Archives of Clinical Neuropsychology 27 (2012) 735–741

Archives of CLINICAL NEUROPSYCHOLOGY

Clinical Utility of Reliable Digit Span in Assessing Effort in Children and Adolescents with Epilepsy

Antoinette J. Welsh[1], H. Allison Bender[2], Lindsay A. Whitman[3], Marsha Vasserman[4], William S. MacAllister[3,*]

- RDS scores showed strong correlations with clinical and cognitive variables, including age of participant and intellectual functioning.
- Overall pass rate of RDS scores at ≤ 6 was low (65%)

Embedded Indicators from CVLT

- Adult studies
  - Dozens of studies across a variety of samples
  - Recognition scores generally most sensitive
- Much less attention in children
  - We've looked at in our mild TBI sample
  - Most recently, N = 411 (aged 8-16 yo)

The Clinical Neuropsychologist, 2013
http://dx.doi.org/10.1080/13854046.2013.858184

Routledge Taylor & Francis Group

Embedded Performance Validity Indicators Within the California Verbal Learning Test, Children's Version

David A. Baker, Amy K. Connery, John W. Kirk, and Michael W. Kirkwood
Department of Physical Medicine & Rehabilitation, University of Colorado Denver School of Medicine and Children's Hospital Colorado, Aurora, CO, USA

Children's Hospital Colorado

Table 3. Summary of logistic regression analysis for five CVLT-C variables predicting adequate versus non-credible effort

CVLT-C Variable	B	*SE*	Wald	*df*	*p*	Odds Ratio	95% CI for OR Lower	95% CI for OR Upper
SDFR	−.591	.339	3.029	1	.082	.554	.285	1.077
SDCR	−.034	.378	.008	1	.929	.967	.461	2.027
LDFR	.488	.361	1.824	1	.177	1.628	.803	3.304
LDCR	.568	.372	2.338	1	.126	1.765	.852	3.657
RD	1.000	.190	27.630	1	.000	2.719	1.873	3.949
Constant	1.932	.173	124.974	1	.000	6.905	NA	NA

SDFR = Short Delay Free Recall; SDCR = Short Delay Cued Recall; LDFR = Long Delay Free Recall; LDCR = Long Delay Cued Recall; RD = Recognition Discriminability.

Table 4. Classification statistics for Recognition Discriminability

Recognition Discriminability *z*-score	Sensitivity %	Specificity %
−0.5	55	91
−1.0	41	97
−1.5	32	98
−2.0	29	99
−2.5	24	99
−3.0	15	100

35

Children's Hospital Colorado

- Again….
- Different results will almost certainly be obtained in lower functioning populations (e.g., those with neurologically or developmentally-based problems)

36

OXFORD
UNIVERSITY PRESS

Archives of Clinical Neuropsychology 30 (2015) 200–206

Archives of CLINICAL NEUROPSYCHOLOGY

Embedded Performance Validity on the CVLT-C for Youth with Neurological Disorders

Brian L. Brooks[1,2,3,4,5,*], Danielle M. Ploetz[1]

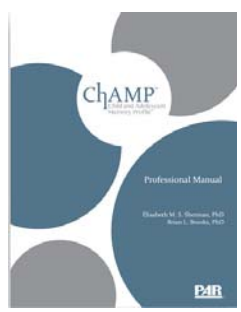
Table 4. Classification statistics for CVLT-C re

CVLT-C recognition discriminability *z* score	Sensitivity	Specificity
0.0	88	59
−0.5	81	62
−1.0	81	68
−1.5	69	79
−2.0	56	84
−2.5	44	88
−3.0	**44**	**90**
−3.5	38	93
−4.0	31	94
−4.5	25	96
−5.0	25	97

37

Child & Adolescent Memory Profile (Sherman & Brooks, 2015)

- First commercially available pediatric test to include embedded indicators
- Brief memory battery with two verbal and two visual subtests
- Subtests contain embedded indicators using three-item forced-choice responding
- Cutoffs based on below chance responding
- 1,200 youth aged 5-21 years, 200 youth with clinical diagnoses, and 45 children in a simulation design study



38

Table 7 Frequency Use of SVTs with Children and Adolescents.

SVT	Never (%)	Rarely (%)	Sometimes (%)	Often (%)	Almost Always (%)	
BASC-2 Validity Indicator	32.2	6.9	12.7	19.2	29.0	48%
BRIEF Validity Indicators	27.2	7.2	12.0	21.0	32.6	54%
MMPI-A Indicators	47.1	15.6	13.4	13.4	10.5	24%
Personality Inventory for Youth Validity Indicators	86.6	4.3	3.9	3.2	2.2	
Trauma Symptom Checklist for Children Validity Indicators	87.3	5.8	4.7	0.7	1.4	

39

Children's Hospital Colorado

BRIEF (and other domain-specific scales)

- No identified independent studies examining faking bad or negativity scales

Article

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

Assessment
1-8
© The Author(s) 2014
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1073191114520626
asm.sagepub.com
SAGE

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

- N = 274; 8-17 year olds administered BASC2
- Primary question
  - What is the relationship between the BASC-2 validity indicators and PVT performance in a sample of real-world pediatric patients?
- Hypothesis
  - Children who failed MSVT would be more likely to have elevations on BASC-2 validity indices, most notably the F index
- Self-Report BASC2 Validity Indicators
  - F Index: designed to assess that a child responded in an inordinately negative fashion or was "faking bad"
  - L index: designed to detect a response set that may be characterized as one of social desirability or "faking good"
  - V index: consists of nonsensical items that may be marked because of carelessness or failure to cooperate or understand questions
  - Response Pattern index: designed to identify forms that may be invalid because of inattention to item content (e.g., N-N-N-N; T-F-T-F-T-F)
  - Consistency index: identifies cases where differing responses given to items usually answered similarly

BASC-2 SRP Validity Scales	MSVT PASS = 224	MSVT FAIL = 50	Significant Test Two-tailed, Fisher's Exact Test
F Scale Within "Caution" or "Extreme Caution" range	4	3	p = .117
F scale Within normal limits	220	47	
Response Pattern Within "Caution" or "Extreme Caution" range	1	1	p = .332
Response Pattern Within normal limits	223	49	
Consistency Scale Within "Caution" or "Extreme Caution" range	10	0	p =.217
Consistency Scale Within normal limits	214	50	
L Scale Within "Caution" or "Extreme Caution" range	14	0	p = .081
L Scale Within normal limits	210	50	
V Scale Within "Caution" or "Extreme Caution" range	2	0	p =1.00
V Scale Within normal limits	222	50	
Any Validity Scale Within "Caution" or "Extreme Caution" range	29	4	p = .471
Any Validity Scale Within normal limits	195	46	

Denver Mild TBI Case Series & BASC2 Conclusions

- First identified study to examine a self-report validity scale in a real-world pediatric sample of noncredible responders
- Vast majority of patients who failed the MSVT provided valid self-report BASC-2 profiles
- Data contrasts with many adult studies demonstrating self-report validity scales strongly associated with PVT performance
- Sole reliance on validity indicators from the BASC-2 (and other child self-report scales?) likely to substantially underestimate the number of patients providing invalid data during neuropsychological evaluation

Table 11 Frequency of Statements to Communicate (Verbally or in Report) Noncredible/Invalid Data.

Statement	Never (%)	Rarely (%)	Sometimes (%)	Often (%)	Almost Always (%)
Test results are invalid	10.6	23.0	40.4	18.1	7.4
Test results indicate inadequate effort to perform well	12.1	12.5	35.8	33.6	6.0
No firm conclusions can be drawn	9.3	16.2	37.4	31.3	6.0
Test results are inconsistent with severity of condition	5.7	12.8	40.8	35.1	5.7
Test results indicate inadequate engagement	15.9	15.9	40.5	23.1	4.5
Test results indicate poor compliance	18.9	20.5	37.9	18.2	4.5
Test results indicate exaggeration or feigning	28.3	31.7	32.8	4.9	2.3
Test results indicate malingering	64.9	28.7	5.7	0.4	0.4

Note. Data are presented in descending order based on "Almost Always".

44

Children’s Hospital Colorado Opening Statement for Parent Feedback in Face of Noncredible Effort (Connery, Baker, Peterson, & Kirkwood)

“Whenever we do these evaluations, we give tests that measure whether children are trying their best to do well in order to make sure the test results are valid. In other words, when a child does not do well on testing, we want to make sure that it is due to an actual weakness rather than to a child not trying his/her best. During today’s evaluation, these tests showed that XXX was not always trying his/her best to do well. What are your thoughts about this? Do you have ideas on why this might have happened?”

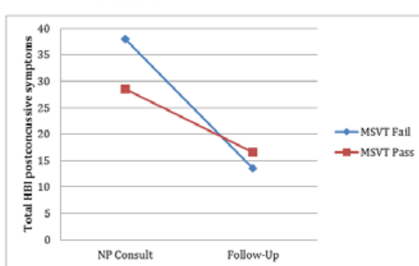
45

Initial Data Regarding Effectiveness of Colorado Feedback Model

- In general, very high rate of satisfaction with neuropsych service in mild TBI clinic (~95%)
  - Kirkwood, Peterson, Connery, & Baker, in submission
- Examination of service in credible vs. noncredible responders (Connery, Peterson, Baker, & Kirkwood, in submission)
  - No difference in caregiver satisfaction rates
  - Actually see greater symptom reduction in noncredible responders

Figure 1.

Self-reported Postconcussive Symptoms.



46

Table 13 Justification for Using or Not Using Validity Testing.

Justification	% of Respondents
Justification for using PVTs	
Independent research supports their utility.	76.5
They are necessary to validate other test results.	68.3
My own experience leads me to believe I need them.	64.9
Practice organizations recommend their use.	50.6
Their use protects examinees.	27.7
Their use protects me from allegations of misconduct.	23.0
Third parties insist on it (e.g., College Boards).	18.1
None—I rarely or never use PVTs in my practice with those under 18 years of age.	9.8
I have additional reasons for using PVTs in my pediatric practice not captured here.	76.8
Justification for not using PVTs	
They are difficult to interpret in very young children (e.g., under 6 years of age).	50.8
They are difficult to interpret in the face of severe cognitive impairment.	38.9
Exaggeration or feigning is usually obvious in a child's general presentation.	18.8
They take too much time.	16.7
Exaggeration or feigning is usually obvious in the pattern of a child's test scores.	13.4
They are difficult to interpret in those under 18 years of age.	11.1
The yield in most cases is not worth the financial cost.	9.6
Clinical cases rarely exaggerate or malinger so they are typically unnecessary in non-forensic settings.	8.4
I have not received adequate training to use them.	7.5
Third parties do not pay for them (e.g., SSI disability).	5.0
Too many genuine patients or claimants are wrongly classified by these tests.	2.5
They are unreliable.	2.1
Identification of exaggeration or feigning might harm the child.	2.1
Identification of exaggeration or feigning might harm the reputation of my practice.	1.7
None—I almost always or always use PVTs in my practice with those under 18 years of age.	30.5
I have additional reasons for not using PVTs in my pediatric practice not captured here.	14.6

47

Rationale for Using PVTs with School-Aged Children and Adolescents

1) Children are capable of deception

2) Noncredible presentations occur consistently in pediatric cognitive assessments

3) We have empirically-backed objective methods to help detect invalid data – why not use?

4) Failure on PVTs has significant implications
- Data interpretation
- Clinical management
- Systemically

Kirkwood (2015). A rationale for including performance validity testing in child and adolescent assessment. In M.W. Kirkwood (Ed.). *Validity Testing in Child and Adolescent Assessment: Evaluating Exaggeration, Feigning, and Noncredible Effort*.

General Pediatric Clinical Case Series

Source	Population	N	Age	PVT	% Noncredible
Donders (2005)	Mixed Neuro	100	6 – 16	TOMM	2%
Carone (2008)	Moderate-Severe Brain Injury	38	(mean: 11.8)	MSVT	5%
MacAllister, Nakhutina, Bender, Karantzoulis, & Carlson (2009)	Epilepsy	60	6 – 17	TOMM	3%
Green et al. (2010)	Mixed Neuro/Dev	380		WMT	5%
Green et al. (2010)	Mixed Neuro/Dev	265		MSVT	3%
Kirk, Harris, Hutaff-Lee, Koelmay, Dinkins, & Kirkwood (2011)	Mixed Neuro/Dev	100	5 – 16	TOMM	4%
Brooks (2012)	Mixed Neuro	100	6 – 19	VSVT	5%
Ploetz, Mosiewicz, Kirkwood, Sherman, & Brooks (2014)	Mixed Neuro	266	5 – 18	TOMM	3%

Pediatric Case Series: Mild TBI

Source	Population	N	Age	PVT	% Noncredible Presentation
Children's Hospital Colorado Kirkwood & Kirk (2010); Kirkwood et al. (2011); Kirkwood et al. (2012); Kirkwood et al. (2013); Baker et al. (2013); Green et al. (2014); Kirk et al. (2014); Kirkwood et al. (2014) Larson et al. (2015)	Mild TBI (clinical)	1000+ total	8 – 17	MSVT + TOMM Rey FIT Various embedded measures	12 – 19%
Araujo et al. (2014)	Mild TBI (clinical)	382	8 – 16	RDS Digit Span	20%

Pediatric Case Series: Independent Setting (Social Security Disability)

Source	Population	N	Age	PVT	% Noncredible Presentation
Chafetz et al. (2007); Chafetz (2008)	Social Security Disability Claimaints (independent)	123	6 –16	TOMM MSVT	48-60% (26-30% PVT chance level or below)

51

Implications of PVT Failure for Interpreting Other Data During a Cognitive Exam

SO WHAT?

- Multiple studies with adults have suggested that PVT performance relates strongly to ability-based tests
  - Green et al., 2001; Constantinou et al., 2005; Green, 2007; Lange et al., 2010; Meyer et al., 2011
  - In these samples (mostly compensation-seeking), ~50% variance in neuropsychological test scores explained by PVT performance (much more variance than explained by brain injury severity, education, age, etc.)
- Up until few years ago, no identified studies in pediatric populations: similar effects?

Children's Hospital Colorado

Psychological Assessment

1040-3590/11/$12.00 DOI: 10.1037/a0024628

The Implications of Symptom Validity Test Failure for Ability-Based Test Performance in a Pediatric Sample

Michael W. Kirkwood
University of Colorado Denver School of Medicine and Children's Hospital Colorado, Aurora, Colorado

Keith Owen Yeates
The Ohio State University and Nationwide Children's Hospital, Columbus, Ohio

Christopher Randolph
Loyola University Medical Center, Maywood, Illinois

John W. Kirk
University of Colorado Denver School of Medicine and Children's Hospital Colorado, Aurora, Colorado

Table 1
Background and Injury Characteristics of All Participants

Participants	$N = 276$
Age (years)	$M = 14.2$, $SD = 2.2$
Grade	$M = 8.3$, $SD = 2.2$
Male	$n = 172$ (62%)
Caucasian	$n = 232$ (84%)
Estimated Full Scale IQ[a]	$M = 103.5$, $SD = 12.6$
Maternal years of education	$M = 15.1$, $SD = 2.2$
Paternal years of education	$M = 15.2$, $SD = 2.6$
Premorbid history of attention-deficit/hyperactivity disorder	$n = 45$ (16%)
Premorbid history of diagnosed learning disability	$n = 29$ (11%)
Premorbid history of special education services	$n = 35$ (13%)
Weeks since injury	$M = 9.7$, $SD = 9.1$; $Mdn = 6.0$
Loss of consciousness	$n = 49$ (18%)
Neuroimaging conducted	$n = 200$ (73%)
Intracranial findings on computed tomography or magnetic resonance imaging for those who underwent neuroimaging	$n = 27$ (14%)
Families in or planning litigation	$n = 22$ (8%)
Families seeking disability compensation	$n = 0$
Participants charged with a crime	$n = 0$

Children's Hospital Colorado

If MSVT measures effort, not ability, two expectations:

1) MSVT performance should be unrelated to demographic, developmental, and injury-related factors in sample

2) MSVT performance should relate to a wide range of tests across the battery – not just those tests tapping memory or related skills (e.g., reading) that seem necessary on surface to complete the MSVT

Support for idea PVT measures effort rather than ability

- No background or injury-related variable differentiated those who passed from those who failed
- PVT results explained ~40% of the variance across the test battery

Table 5
Descriptive Statistics and Comparisons Between Medical Symptom Validity Test Pass and Fail Groups on Ability-Based Tests

Test	Pass			Fail				
	n	*M*	*SD*	*n*	*M*	*SD*	*p*	*d*
WASI								
Estimated IQ	215	105.5	11.6	48	94.5	13.4	<.001**	0.9
Vocabulary T score	215	53.6	8.6	48	50.7	10.9	.045	0.3
Matrix Reasoning T score	215	52.4	7.2	50	41.0	10.6	<.001**	1.4
CVLT–C								
Total Learning Trials 1–5 T score	186	53.0	8.4	40	46.6	11.4	.002**	0.7
Long Delay Free Recall *z* score	186	0.34	0.8	40	−0.48	1.3	<.001**	0.9
Recognition Discriminability *z* score	186	0.18	0.6	40	−1.29	1.8	<.001**	1.6
WISC–IV								
Digit Span scaled score	224	9.9	2.9	51	6.4	3.2	<.001**	1.2
Coding scaled score	207	9.7	5.3	45	6.4	3.1	<.001**	0.6
Grooved Pegboard								
Dominant hand *z* score	213	−0.25	1.4	45	−1.7	2.5	.001**	0.9
Nondominant hand *z* score	215	−0.41	1.5	45	−1.6	2.2	.001**	0.7
Woodcock–Johnson III								
Letter–Word Identification standard score	191	100.2	9.7	45	97.0	22.0	.347	0.3
Automatized Sequencing (time in seconds)								
Alphabet	216	5.6	6.1	50	11.4	10.9	.001**	0.8
Counting 1 to 20	172	4.7	1.4	44	9.6	12.5	.013	0.9
Days of week	209	2.5	1.2	47	5.4	5.1	<.001**	1.2
Months of year	214	6.1	4.4	47	12.0	6.8	<.001**	1.2

Implications of PVT Failure for Interpreting Symptom Report

- Several studies with adults have also suggested that performance on PVTs has significant effect on postconcussive symptom report after mild TBI
  - Lange et al. (2010)
  - Iverson et al. (2010)
  - Tsanadis et al. (2008)
- No identified studies in pediatric populations: similar effects?
- In Denver mild TBI series, children failing MSVT reported significantly more “postconcussive symptoms” than those who pass MSVT ($p < .001$; $d = 1.1$)
  - Kirkwood, Peterson, Connery, Baker, & Grubenhoff (2014)

PEDIATRICS

OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

Postconcussive Symptom Exaggeration After Pediatric Mild Traumatic Brain Injury

Michael W. Kirkwood, Robin L. Peterson, Amy K. Connery, David A. Baker and Joseph A. Grubenhoff

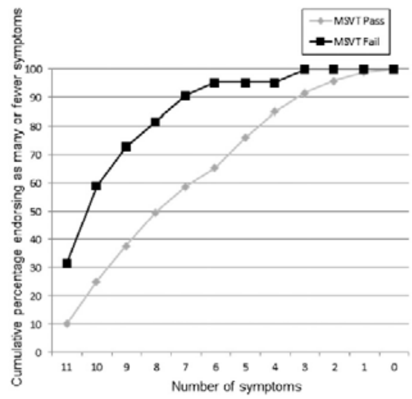


FIGURE 1
Cumulative percentage of total number of postconcussive symptoms endorsed for MSVT Pass and Fail groups.

57

Implications of PVT Failure for Broader Systems

- As one example….
- Social Security Administration
  - In 2011, for malingered mental disorders in adults, estimated cost to SSA was $20.02 billion
    - Chafetz & Underhill (2013). Estimated costs of malingered disability in Social Security Disability examinations. *ACN, 22,* 1-14.
  - In 2011, for malingered mental disorders in children, estimated cost to SSA was $2.13 billion
    - Chafetz (2015). Disability: SSI exams for children. In M.W. Kirkwood (Ed.). *Validity Testing in Child and Adolescent Assessment: Evaluating Exaggeration, Feigning, and Noncredible Effort*.
- Given that many pediatric providers do not routinely use PVTs, likely an underestimate when collectively consider governmental, legal, healthcare, and educational costs

Michael Kirkwood, Ph.D., ABPP/CN

Children's Hospital Colorado
Department of Rehabilitation Medicine

Michael.Kirkwood@childrenscolorado.org
720-777-6193