The Importance of Symptom Validity Testing with Children: WMT and MSVT

Lloyd Flaro Ed.D., Paul Green Ph.D. & Nina Blaskewitz Dipl.Psych.

Abstract: It is almost self-evident that cognitive test results will be unreliable and misleading if children do not make a full effort on testing. Nevertheless, objective tests of effort have not typically been used with children to determine whether test results are valid or not. Four cases are presented in which children's intelligence test scores greatly underestimated their actual intelligence, owing to poor effort that sometimes went undetected. Selected effort tests for use with children are discussed. Objective testing of effort in children is recommended to avoid misinterpreting invalid test data, which is why the use of effort tests is now standard practice in forensic neuropsychology (Iverson, 2006).

Correspondence should be addressed to Dr. Lloyd Flaro in English (201, 17010 103 Ave., Edmonton, Alberta, Canada, T5S 1K7) or Nina Blaskewitz in German (Buchenweg 19, 66265 Holz, Deutschland).

Introduction: All psychological testing with children is an attempt to sample specific behaviors when a child is given specific tasks to complete. Psychological testing demands adherence to standardized administration procedures to ensure the reliability and validity of test results (Sattler, 1998). Standardization permits the clinician to establish procedural consistency and control over the administration of psychological tasks by holding constant the testing protocol. The ideal procedural methodology makes sure that, as far as possible, children are subjected to the same tasks in the same manner (Anastasi, 1988; Kamphaus, 1993; Kaufman, 1994). An integral part of this procedural methodology is adherence to the recommended guidelines for establishing rapport and the promotion of optimal effort levels across psychological and neuropsychological tasks.

An attempt is made by the psychological administrator to elicit the best performance from the child during the testing. The tester tries to engage the child's full cooperation and to encourage the child to pay close attention to the task demands and take the testing seriously (Reitan & Wolfson, 1997). The tester tries to establish environmental conditions that minimize distractions, increase active participation, stimulate interest in the tasks and minimize anxiety and fear responses (Kaufman, 1994; Prifitera & Saklofske, 1998, 2005). This requires the ability of the examiner to establish rapport with the child (Kaufman & Lichtenberger, 2000) and the ability to make sure that the child works hard during the psychological or neuropsychological evaluation.

Despite a general acceptance of the need for optimal conditions to encourage the child's best effort on testing, little research has been produced that details how to measure effort levels in children objectively and how to identify when effort is suboptimal. A large percentage of clinicians use clinical judgment to determine the degree of effort put forth on the tasks by children. However, research has suggested that clinicians' judgment of effort levels with children are inaccurate, deficient and clinically faulty (Dawes, Faust & Meehl, 1989; Oldershaw & Bagby, 1997). Faust and Hart (1988) added to this research by demonstrating that when children were coached into demonstrating neurocognitive impairment, most clinicians failed to detect the malingering process. Faust, Hart, Guilmette and Arkes (1988) replicated the study with

adolescents and found the same results. Clinical judgment may not be enough to identify invalid data owing to poor effort.

Poor performance on psychological and neuropsychological tasks may result from actual cognitive impairment and define a clinical picture accurately representing the child's overall ability profile. On the other hand, children who deliberately put forth poor effort will produce test results that are invalid and unreliable. Some clinicians discern poor effort by examining consistency levels across task performance (Kaufman, 1994; Reitan & Wolfson, 1997). Others simply assume that children have no vested interest in performing poorly in any psychological or neuropsychological evaluation. The assumption that children will exert consistent and optimal levels of effort in the testing has been challenged by a number of researchers (Faust & Hart, 1988; McCaffrey & Lynch, 1992; Rogers, 1997). In the area of ADHD, it is known that children with this clinical disorder often perform poorly on a variety of academic and psychological tasks and demonstrate significant motivation difficulties, even though there is evidence of adequate cognitive capacities (Barkley, 1999; Douglas, 1983, Flaro & Green, 2000, 2003).

The lack of objective measures for assessing children's effort levels in psychological and neuropsychological testing primarily results from an emphasis on determining the antecedent causes of poor effort. This may be a viable theoretical position but the fact of the matter is that poor effort contaminates the validity and the reliability of the test results, which sometimes leads to serious misinterpretations and misdiagnoses. In the literature little attention has been paid to the consequences of poor effort, which include inconsistent task performance, inconsistent test results over time, overestimation of impairment and misdiagnosis. To illustrate the importance of effort testing in children, several case studies will be presented.

Case 1: Andrew's FSIQ increased by 37 points

Andrew was a nine year old boy diagnosed with a reading disability and low average intellectual abilities. His school teacher and mother referred Andrew for a psychological evaluation because they did not believe that Andrew was functioning as poorly as reported in a

previous psycho-educational evaluation completed by another psychologist in January of 2002. A reading specialist reported that Andrew was a bright child with a high level of ability to process information and to answer questions. His ability to read was much higher than predicted by his intellectual abilities.

In the previous psychoeducational evaluation, Andrew's full scale intelligence (FSIQ) was found to be 85. However, when seen by the first author for a psychological evaluation in February of 2003, his FSIQ was 122 (superior range on the WISC-IV). He was reading at a grade seven level, consistent with the reading specialist's estimation. The difference between FSIQ values of 85 and 122 was obtained despite using the same test and it could not reasonably be accounted for by practice effects. The difference was simply too great.

When Andrew was questioned about his previous assessment, he spontaneously said "*The man was cruel and mean. He did not talk to me very much*". Andrew further explained " *I was mad and did not do my best*" (poor effort). This is a case in which the initial examiner violated two important principles of testing; gaining rapport and promoting optimal effort. The third error was failing to measure effort objectively. The consequence of the inaccurate initial FSIQ score was damaging because Andrew was diagnosed as a "slow learner with severe reading disabilities" and he was placed in a special program, aimed at working with him "at his level". As expected, Andrew was quite bored and disgruntled with the whole situation and his education underwent a set-back because of a misdiagnosis.

Case 2: Eric's FSIQ dropped before a court appearance

Eric was a twelve year old Métis boy referred for a neuropsychological assessment by his psychiatrist. He had recently been in secure treatment for adolescents diagnosed with severe behavioral problems and was awaiting trial for criminal charges. The psychiatrist described him as being violent and aggressive, socially inappropriate and heavily involved with drugs and alcohol. He lacked impulse control and had a need for high sensation-seeking activities, leading him into many dangerous situations. The psychiatrist wanted to know if his maladaptive

behavioral pattern reflected underlying neuropsychological impairment or was the result of developmental personality dysfunction and social learning.

During the testing, Eric presented in a defended and guarded manner and made it clear that he would complete the testing on his own terms. He presented as an adolescent with severe conduct disorder and emerging antisocial tendencies. His effort during this assessment was variable and questionable; he passed one objective measure of effort (Word Memory Test; IR=100; DR=100; CNS=100; MC=95; PA=100; FR=67.5) and failed the other test (Medical Symptom Validity Test; IR=90; DR=80; CNS=70; PA=90; FR=40). In addition to failing one effort measure, clinical observation of his test performance indicated inconsistent results across a number of tasks.

During the assessment, Eric reported that he had to stand trial for his aggressive behavior (assault on another male). He beat him severely with a tire iron and demonstrated little remorse for his actions. Indeed, he blamed this person for the assault because he had acted in a certain manner which made Eric angry. Eric was concerned that this testing might keep him in secure treatment, whereas he wanted to be placed in a group treatment home with more freedom. When challenged on this placement, he laughed and admitted that he could go absent without leave easily at this place, allowing him ready access to drugs and alcohol.

Of particular note was the significant difference in the results of intellectual assessments completed over time. In the current assessment, his Full Scale IQ was 77, whereas in a previous assessment a year earlier his general intellectual functioning was average (FSIQ of 102). Over one year, his intellectual scores had declined by 25 IQ points, a finding that made no sense given his clinical background. In addition, in the current assessment his verbal IQ was 83 while in the previous assessment it was 110. His nonverbal ability was 76 compared with 93 in the previous assessment. Clearly, this case represented an antisocial adolescent who was trying to look more impaired than he really was at the time. It would be hard to defend interpreting his

neuropsychological results as valid and reliable, given his obvious attempt at manipulation of the situation and the examiner.

When Eric was confronted with the discrepancies in his results, he simply stared at the examiner and laughed. Asked to explain the reasons for poor effort, Eric said that he liked playing games with people. He was impressed that the examiner was able to catch him at his own game. This comment, in itself, suggested intelligence not demonstrated by his results in this assessment. When the results were brought to the attention of the psychiatrist, he was not surprised and suggested that Eric had already met the criteria for antisocial personality disorder, even though he was not yet eighteen years old. He predicted that, within the year, Eric would probably be in jail for antisocial and criminal activities. The psychiatrist suggested that Eric would probably react toward others in a violent and aggressive way and seriously hurt them. This was very consistent with the first author's impressions of this adolescent.

Case 3: Melissa's reading level and FSIQ increased significantly over time

Melissa was a sixteen year old girl, from a high achieving, wealthy professional family, who was referred for a neuropsychological evaluation by a friend of the family, Dr. Y, who requested assessment of a wide range of cognitive abilities. Dr. Y was concerned that there might be significant reading difficulties, holding back Melissa's progress in school. He noted his perception of personality differences between Melissa and her older, highly academic sister.

During a clinical interview with the second author, Melissa said that she really did not see the importance of formal academic studies. She indicated that she had problems with reading, saying that she read slower than her friends and that she did not like studying because it was too difficult. As a result of her reading difficulties, she claimed that she had become tired of school, which required too much effort.

Melissa was administered a number of effort measures and she achieved normal range performance on all of them. Hence, we assume that her neuropsychological results were valid and

reliable. On a measure of intellectual functioning (Multidimensional Aptitude Battery), her results were in the superior range and at the 93rd percentile. Reading skills on the Wide Range Achievement Test-3 were in the average range (53rd percentile). Listening abilities, memory functioning, abstract abilities, attention, visual search and alternating mental set were all within normal to above average ranges.

Assessment of personality functioning with the MMPI-2 indicated the presence of antisocial tendencies. A high score on scale 4 suggested difficulty incorporating the values and standards of society and suggested that she was prone to engaging in a wide range of antisocial and conduct-disordered behavior. It would be expected that Melissa would be rebellious toward authority figures. Moreover, the personality test results characterized Melissa as probably being impulsive and having significant difficulty with delaying gratification.

When the current results were compared with those from previous testing, some important discrepancies emerged. In the previous assessment her, intelligence was only average (FSIQ=94), whereas it was now in the superior range (FSIQ=122). No objective tests of effort were used in the previous psychological evaluation. The most likely explanation for her poor showing in the initial assessment was that she was not putting forth her best effort, which would be consistent with her antisocial tendencies and her expressed attitude of boredom and lack of motivation in school. In the current assessment, in contrast, she passed all effort measures.

In a similar vein, her reading comprehension skills in the previous assessment were noted to be below the normal mean and at only the 23rd percentile. In this assessment, her reading comprehension skills were found to be at the 96th percentile using the Woodcock-Johnson Passage Comprehension subtest. This represents another discrepancy between test results produced over time. Melissa's personality features and the presence of antisocial tendencies were thought to be contributing to significant differences in test results over time. It was also noted that there was a strong possibility of a mood disturbance. Her mood was slightly elevated when found

to be making a full effort but she had previously been through at least one episode of major depression.

Case 4: Paul's IQ increased 25 points in one day

Paul was a seven year old Caucasian boy, referred by his parents for a comprehensive psychological assessment. The parents were concerned about his social impairment, difficulty adapting to any sort of change, overreaction to minor events, growing dependence on his parents and unwillingness to socialize with other children. The parents were concerned that he might be demonstrating the signs of emotional and behavioral disturbance seen in children diagnosed with Asperger's Syndrome. On the first morning of the testing, Paul presented as shy and timid. He lacked in social engagement, social initiative and social reciprocity. He failed to establish eye contact throughout the testing. He perseverated on topics of interest, which made it difficult to bring him back to task. On a number of initial tasks administered to him he gave up easily even though he was informed of the importance of putting forth his best effort on the tasks administered to him.

He was administered the Wechsler Intelligence Scale for Children-IV. His Full Scale IQ was 88 (21^{st} percentile). While observing him complete the tasks in the WISC-IV it became apparent that he was not putting forth his best effort. This was confirmed by his poor performance on one effort measure (Medical Symptom Validity Test: IR=90; DR=80; CNS=70; PA=90; FR=40). His poor performance on the WISC-IV was brought to the attention of the parents, who requested that he be retested after they had a talk with him regarding his motivation and effort. On the next morning, he was given the Wechsler Abbreviated Scale of Intelligence and his Full Scale IQ was 116 (86^{th} percentile), an increase of 28 points, compared with the day before.

Of particular interest was the difference of twenty-five points between his previous Verbal IQ of 79 and his current WASI Verbal IQ of 104. His nonverbal intellectual score increased from 107 to 122. Paul's effort level appeared to be good in this assessment and it was confirmed by his

normal performance on a second effort measure, the Nonverbal Medical Symptom Validity Test (scores in % correct: IR=100; DR=95; CNS=95; DRA=80; DRV=90; PA=100; FR=50).

If a seven year old boy can demonstrate significant differences in intellectual ability in a day and a half of testing, what does this say about the reliability and the validity of the test results? It was crystal clear that his effort level was responsible for his poor results on the WISC-IV and that, after a pep talk by his parents, his effort level increased and he showed positive changes in his cognitive test scores.

Effort tests for use with children

All four cases discussed above demonstrate the importance and the necessity of using objective measures of effort in psychological and neuropsychological evaluations. Without these effort measures, we must rely on our clinical judgment for the analysis of inconsistent performance. In two of these cases, clinical judgment failed and invalid data were interpreted as valid in the initial assessments. Poor effort has serious consequences, including inconsistent task performance, underestimation of the child's true abilities and misdiagnosis.

Comments have traditionally been made about effort in nearly all assessment reports, reflecting the widespread belief that the effort applied to testing is of critical importance to the validity of the test results. It is surprising, therefore, that until recently there were no objective ways to measure effort levels in children. Over the last decade, however, several researchers have explored the possibility of testing children using symptom validity tests originally devised for use with adults, such as the WMT (Flaro & Green, 2001, Green & Astner, 1995, Green, Allen and Astner, 1996, Green, 2003) and the TOMM (Donders, 2005). New effort measures have been created for use with children, such as the MSVT (Medical Symptom Validity Test, Green, 2004) and its nonverbal equivalent the NV-MSVT (Green, 2006). The WMT and the MSVT are available in several languages, including German, French, Spanish and English. The NV-MSVT

requires no reading skills and consists of memory testing using colored images drawn by an artist and presented on a computer screen.

Initial research showed that the Word Memory Test was a viable psychological instrument that could be used with children (Flaro & Green, 2000). In an extension of the latter study, 135 children between the ages of seven and eighteen years, as a group, achieved results on the WMT effort subtests similar to those from parents seeking custody of their children (Green and Flaro, 2003). In this study, most children diagnosed with Fetal Alcohol Syndrome, Schizophrenia, Bipolar Disorder, ADHD, Conduct Disorder and Learning Disabilities demonstrated no difficulty exceeding the adult cut-offs for good effort on the Word Memory Test, as long as they were seven years of age or older and with a grade three reading level or higher. In children with less than a grade three reading level, the oral version of the WMT was recommended.

Recent research on the Word Memory Test continues to support this instrument as a potentially valuable objective measure of effort in children. The first author (LF) has administered the Word Memory Test to two hundred and twenty-six children between the ages of seven and eighteen. As seen in Figure 1, on the effort measures (IR, DR, CNS) the children achieved results between 93% and 96%, values comparable to adult groups (Figure 2).

It is an astonishing fact that, in general, far more adults than children fail the WMT. Many adults tested with the WMT have an incentive to appear more impaired than they are because they are seeking compensation for disability. In over 2,000 adults tested by the second author and by Dr. Roger Gervais (personal communication), 30% of cases failed the WMT effort subtests, compared with only 11% of 263 children tested clinically.

In a recent analysis of data from 116 parents seeking custody of children, the first author found that only two cases failed the WMT effort subtests. This is a very important group because it consists of adults who had a positive incentive to appear competent. The Court and the Department of Social Services relied on the results of the assessment to determine whether these

parents were to be given custody of their children. In 65% of cases, custody of their children was denied. This group contained many people with very significant cognitive impairment, which was evident, for example, in the fact that 60% of cases made more than 60 errors on the Category Test and 20% had a FSIQ of less than 80. Nevertheless, 98.2% of cases passed the WMT effort subtests (i.e. scored more than 82.5% correct on immediate recognition, delayed recognition and consistency of performance across the two subtests). Those judged unfit to be parents were of lower than average intelligence (mean FSIQ=86, SD 12) but their mean scores on the WMT effort subtests were almost perfect, at 98% correct (SD 4) on immediate recognition, 98% correct (SD 3) on delayed recognition and 97% (SD 4) on consistency. In this sample of adults, the existence of false positive results (i.e. cases who failed WMT despite making a good effort) was zero. Two parents failed the WMT but they both admitted that they had made a poor effort on testing. They had changed their minds in the course of a drawn out custody battle and now did not want their children returned to their care. A year later, one of these parents returned for testing and passed the WMT.

Insert tables 1 & 2 here

An even easier effort measure than the WMT was developed to assess effort in children and the preliminary results of this test (the Medical Symptom Validity Test, MSVT, Green, 2004) are most promising. In the MSVT version, the child must learn ten word pairs, instead of 20 in the WMT. Having only half as many word pairs makes the test even easier than the WMT. The word list is presented twice on the computer. Also, each word pair conceptually reflects a single idea (e.g. eye-ball), whereas the pairs represent two concepts in the WMT (e.g. tree-lake). On the immediate recognition (IR) trial, the child is shown two words, one from the target list and a new foil word. The child is asked to select the target word (i.e. the word seen before). After a ten minute delay, the child is asked to perform a similar delayed recognition (DR) test, followed by a paired associate task (e.g. “What went with the word eye?”). Finally, on the free recall task, the child is asked to recall as many words as possible from the original list.

To establish normative data on the MSVT, unselected school children in grades two to seven were given the computerized MSVT. Most were healthy children with no learning disability but some were in special classes. The results from all these children combined are shown in table 2. The median scores on the IR and DR subtests were 100% correct at most age levels. The mean correct was 99% (SD 3) on both IR and DR in the sample as a whole. Only two children out of 96 scored 85% or lower on the immediate or delayed recognition subtests.

In grade three, the mean MSVT score on IR and DR was 99% correct. These results demonstrate how easy the MSVT is to pass, where the cutoff is set at 85% for IR and DR subtests. The same children performed less well on the Paired Associate and Free Recall subtests and there were clear age effect on these subtests, as expected. Further data were collected from children seen in clinical practice over a number of years. The data provided in table 3 indicate that children with diagnoses, including Fetal Alcohol Spectrum Disorder, Childhood Schizophrenia and Autism easily passed the effort measures.

Insert table 3 here

Conclusions: We know that a child's level of effort during the assessment will determine the validity and the reliability of the test data. When children fail to put forth their best effort, we cannot be sure of their actual ability profile. As demonstrated in the four cases presented in this article, failing to identify less than optimal effort levels can have detrimental effects on the child in terms of diagnosis, placement, programming, treatment intervention and future educational opportunities. Without some way to measure children's effort levels objectively during an assessment, our conclusions must be considered tentative at best. At the worst, they must be considered inaccurate and invalid. In the emerging field of symptom validity testing in children, we now have a number of objective effort measures, including the WMT, the MSVT and the nonverbal MSVT, which provide us with the means to detect objectively when effort is below the level needed to obtain valid test results. In future, it is likely that any psychologist or neuropsychologist working with children will be required to employ objective effort measures as

an integral part of the test battery, just as it is now the accepted practice to incorporate such tests in adult neuropsychological assessments (Bush, Ruff, et al., 2005, Iverson, 2006).

Table 1: Mean scores on WMT subtests in 264 children tested clinically, aged 7 to 18 years

	Primary effort subtests			**Secondary effort subtests**		**Memory subtest**		
	WMT IR	WMT DR	Consistency	Multiple Choice	Paired Associate Recall	Free Recall	Age	FSIQ
Mean	**96%**	**96%**	**93%**	**86%**	**83%**	**46%**	**13%**	**89**
SD	7	8	9	19	19	18	2.8	15
Median	97.5%	97.5%	95%	95%	90%	48%	13	89
Minimum	45	30	20	5	0	0	7	53
Maximum	100	100	100	100	100	95	18	139
% of group scoring >82.5%	95%	94%	90%	76%	65%	1%	-	-

Table 2: Mean, median, standard deviation and fifth percentile values for MSVT subtests in unselected school children from age 8 to 14 years.

Age (n failing MSVT)		N	Median	Mean	Std Deviation	5th%ile
8	IR	5	**100**	94	11	75
(1	DR	5	**95**	92	10	75
failed	CNS	5	**95**	88	16	60
MSVT)	PA	5	**70**	60	39	10
	FR	5	**50**	47	21	15
9	IR	18	**100**	99	2	90
(1	DR	18	**100**	99	4	85
failed	CNS	18	**100**	99	4	85
MSVT)	PA	18	**100**	89	26	10
	FR	18	**68**	67	16	40
10	IR	34	**100**	99	2	95
(0	DR	34	**100**	99	3	90
failed	CNS	34	**100**	99	3	90
MSVT)	PA	34	**100**	98	5	80
	FR	34	**75**	75	12	55
11	IR	5	**100**	98	5	90
(0	DR	5	**100**	100	-	100
failed	CNS	5	**100**	98	5	90
MSVT)	PA	5	**100**	96	9	80
	FR	5	**80**	84	11	70
12	IR	13	**100**	99	2	95
(0	DR	13	**100**	100	-	100
failed	CNS	13	**100**	100	-	95
MSVT)	PA	13	**100**	94	17	40
	FR	13	**80**	79	21	20
13	IR	9	**100**	100	-	100
(0	DR	9	**100**	100	-	100
failed	CNS	9	**100**	100	-	100
MSVT)	PA	9	**100**	100	-	100
	FR	9	**90**	85	9	70
14	IR	12	**100**	99	2	95
(0	DR	12	**100**	100	-	100
failed	CNS	12	**100**	99	2	95
MSVT)	PA	12	**100**	98	4	90
	FR	12	**88**	85	8	70

Table 3: Mean MSVT scores from 148 children tested clinically in English by Lloyd Flaro (Canada) and 20 children tested in German by Nina Blaskewitz (Germany)

	Clinical testing in Canada			Healthy children in Germany	
	Mean	**SD**	**Median**	**Mean**	**SD**
MSVT Immediate Recognition	98%	5	100%	99.5%	1.5
MSVT Delayed Recognition	97%	5	100%	99.8%	1.1
MSVT consistency IR-DR	97%	6	100%	99.3%	1.8
MSVT Paired Associate Recall	93%	16	100%	80%	25.8
MSVT Free Recall	66%	20	70%	64.5%	21.5
AGE	12	3	12	9	-
FSIQ	87	15	85	-	-

References

Anastasi, A. (1988). *Psychological Testing* (6th ed.). New York: Macmillan.

Bush, S., Ruff, R., Troster, A., Barth, J., Koffler, S., Pliskin, N., Reynolds, C. & Silver, C. (2005) Symptom Validity Assessment: Practice issues and medical necessity (NAN Policy and Planning Committee). *Archives of Clinical Neuropsychology,* 20, 419-426.

Kamphaus, R. W. (1993). Clinical Assessment of children's intelligence. Boston: Allyn & Bacon.

Kaufman, A. S. (1994). *Intelligent testing with the WISC-111*. New York: Wiley.

Reitan, R. M & Wolfson, D (1997). *Neuropsychological Evaluation of Older Children*. South Tucson: Arizona, Neuropsychology Press.

Prifitera & Saklofske (1998). WISC-111: *Clinical Use and Interpretation*. San Diego: California, Academic Press

Prifitera, A., Saklofske, D. H., & Weiss, L. G. (2005). *WISC-IV: Clinical Use and Interpretation*. San Diego: California, Elsevier Academic Press.

Kaufman, A. S. & Lichtenberger, E. O (2000). *Essentials of WISC-111 and WPPSI-R assessment*. New York: Wiley.

Dawes, R. M., Faust, D. & Meehl, P. (1989). Clinical versus actuarial judgment. *Science*, 243, 1668-1674

Faust, D., Hart, K., Guilmette, T. J. & Arkes, H. R. (1988a). Neuropsychologists' capacity to detect adolescent malingerers. *Professional Psychology: Research and Practice*, 19, 508-515.

Faust, D., Hart, K., Guilmette, T. J. & Arkes, H. R. (1988b). Pediatric Malingering: the capacity of children to fake believable deficits on neuropsychological testing. *Journal of Consulting and Clinical Psychology*, 56, 578-582.

Iverson, G., (2006) Ethical issues associated with the assessment of exaggeration, poor effort and malingering. *Applied Neuropsychology*, 13, 2, 77-90

McAffrey, R. J. & Lynch, J. K. (1992). A methodological review of "Method Skeptic" reports. *Neuropsychological Review*, 3, 193-214.

Rogers, R. (1997). Current Status of Clinical Methods. In R. Rogers (Ed.). *Clinical assessment of malingering and deception* (pp.. 373-397). New York: Guildford Press.

Barkley, R. A. (1999). Theories of attention-deficit/hyperactivity disorder. In H. C. Quay & A. E. Hogan (Eds.), Handbook of disruptive behavior disorders (pp. 295-313). New York: Kluver Academic Publishers Press/Plenum.

Douglas, V. I. (1983). Attentional and cognitive problems. In M. Rutter (Ed.). *Developmental neuropsychiatry* (pp. 49-66). New York: Guildford Press.

Donders, J. (2005) Performance on the Test of Memory Malingering in a mixed pediatric sample. *Child Neuropsychology*, 11, 2, 221-227

Flaro, L. & Green, P. (2000). Symptom validity test results with children: CARB and WMT. *Archives of Clinical Neuropsychology*. 15, 840.

Green, P., (2004) The Medical Symptom Validity Test: Test Manual and MS Windows computer program. Green's Publishing, Edmonton.

Green, P., (2006) The Nonverbal Medical Symptom Validity Test: MS Windows computer program. Green's Publishing, Edmonton.

Green, P. & Astner, K. (1995). Test Manual for the Oral Word Memory Test, Cognisyst, N.C. (Out of print but available from Green's Publishing, Edmonton).

Green, P., Allen, L. & Astner, K. (1996). Word Memory Test manual (DOS version), Cognisyst, NC (Out of print).

Green, P. & Flaro, L. (2003), Word Memory Test performance in children. *Child Neuropsycholog*y, 9, 3, 189-207.

Green, P., Iverson, G. & Allen, L. (1999). Detecting Malingering in Head Injury Litigation with the Word Memory Test. *Brain Injury*, 13, 813-819.

Green, P., Lees-Haley, P. & Allen (2002). The Word Memory Test and the Validity of Neuropsychological Test Scores. *Forensic Neuropsychology*, 2, 97-124.

Iverson, G., Green, P. & Gervais, R. (1999). Using the Word Memory Test to detect biased responding in head injury litigants. *Journal of Cognitive Rehabilitation*, 17, 4.8.

Sattler, J. M. (1988). *Assessment of Children* (3rd Ed). San Diego, California

McCann, T. J. (1998). Malingering and deception in adolescents: Assessing credibility in clinical and forensic settings. Washington, DC: American Psychological Association.

Oldershaw, L. & Bagby, M. (1997). Children and Deception. In R. Rogers (Ed.). *Clinical assessment of malingering and deception* (pp. 153-166). New York: Guildford Press.