

Effort-testing In Children Undergoing Psycho-Educational Assessment Using The Medical Symptom Validity Test

Helene Flamand
University of Alberta

Abstract

Suboptimal effort affects the reliability and validity of clinical data. Effort measures of a sample of French Canadian schoolchildren who had undergone a psycho-educational assessment were compared with eight other schoolchildren samples presumed to have exerted good effort, using the Medical Symptom Validity Test (MSVT) and the Reliable Digit Span (RDS). The study focused, more specifically, on the MSVT effort measures, Immediate Recognition (IR), Delayed Recognition (DR) and Consistency (CNS) because they are relatively insensitive to impairments. Overall, the majority of children, with and without impairment passed the MSVT effort measures with a fail rate of 13.1% for MSVT and 36.1% for RDS. The correlation between failure of each measure was only moderately significant. The failure rate on the MSVT dropped to 6.1% after removal of participants who read below a grade three level and the correlation between failure on the RDS and failure on MSVT became insignificant. This is consistent with the literature documenting the use of MSVT and WMT with adults, with and without impairments, and supports the assertion that insufficient effort is likely the only explanation for failure on MSVT when the participants have the required reading level [8]. Green suggested considering evidence of a Genuine Memory Impairment Profile (GMIP) in failing protocols (i.e., when the average of IR, DR, and CNS exceeds the average of PA and FR by at least 20 points). In this study, all but one child who failed MSVT demonstrated a GMIP profile.

1. Introduction

In the field of educational psychology, professionals rely heavily on direct observation of children's behavior during the psycho-educational evaluation process in order to determine if the children's performance on psycho-educational measures is truly a valid representation of their current abilities. While it is easy for persons administering tests to identify overtly uncooperative behavior, such as frequently responding with "I don't know," missing random items in a test comprised of items of increasing difficulty, and easily giving up on more difficult items, they find it far more difficult to identify suboptimal effort of investment in the task [16]. Suboptimal effort of investment has direct

implications for the reliability of the test scores and the validity of any interpretations based on test scores or other clinical data. Not satisfied with relying on observation only, neuropsychologists have developed measures to assess the validity of the assessment.

According to Flaro, Green, and Blaskewitz, the lack of objective measures for assessing the level of children's efforts in psychological and neuropsychological testing results primarily from the emphasis in the field of psychology on determining the antecedent causes of poor effort instead of actually measuring the child's effort at the time of testing [6]. According to these authors, poor effort at the time of assessment, which may be identified by inconsistent task performance, contaminates the validity and the reliability of tests results and leads to inaccurate estimation of impairment and to misdiagnosis [6].

This current study compared effort measures of a sample of French Canadian schoolchildren with those of other samples of schoolchildren presumed to have exerted good effort. This analysis was accomplished by comparing scores on two different effort measures, the MSVT and the RDS, obtained from a clinical sample of French Canadian children undergoing a psycho-educational assessment and other available samples of children who had undergone similar procedures.

2. Problem Statement

Effort tests typically involve simple tasks that can be accomplished with great ease, except when significant cognitive impairment exists. The Medical Symptom Validity Test (MSVT) is the symptom validity test used in this study. The MSVT is an automated memory-screening test with integral effort measures.

According to Green, good-effort measures should be as insensitive as possible to impairments [8]. The more that a measure is affected by variations in ability, the less sensitive it is as a measure of effort. Passing the MSVT requires minimal ability and most people making an effort will easily pass the effort subtests [15]. When a person without significant cognitive impairment—adult or child—fails an effort test as part of a comprehensive, neuropsychological assessment, the reliability and validity of the test results are brought into question.

The Reliable Digit Span was originally based on raw scores derived from the Wechsler Adult Intelligence Scale Digit Span subtest [10]. RDS is calculated by summing the longest string of digits repeated without error over two trials under both forward and backward conditions. For example, a person who passes both trials of three digits forward and both trials of two digits backwards but fails on a trial of four digits forward and three digits backwards would earn an RDS score of 5. The conventional cut-off score for adults is 7 or below, which would indicate poor effort [10]. There is no agreed upon cut-off score for children, although some researchers have used the adult cut-off [3]. Only two studies have documented use of RDS with children: Chafetz, Abrahams, and Kohlmaier used the procedure as one of the variables in a model they designed to identify malingering in children and adults; Hargrave, Kirkwood, and Kirk examined sensitivity and specificity RDS with a sample of 275 mild-TBI clinical patients aged 8 to 16 [5] [11]. While a cut-off score of 7 or less produced an unacceptable high false positive rate (32%), the authors found that an RDS cut-off of 5 or less resulted in 51% sensitivity and 91% specificity, comparable with the results in adult studies. Hargrave et al. concluded that scores from Wechsler Digit Span may be only moderately sensitive but are likely to be useful in identifying suboptimal effort in older children and adolescents, at least in those who are relatively high functioning [11].

Although the use of some effort tests such as the Word Memory Test and the Victoria Symptom Validity Test is reported with college students undergoing psycho-educational assessment, the use of the MSVT with children has not yet been widely examined, and certainly not in the context of a psycho-educational setting [8]. In addition, the use of the RDS procedure as a measure of effort in children has been little documented. Available data suggest that the cut-off scores used with adults are not appropriate for children: Blaskewitz, Merten, and Kathman used many effort measures, including the MSVT, and found the majority (59%) of children obtaining full-effort scores on other measures failed RDS with the traditional cut-off score of 7 [3]. They concluded that RDS results appear to be closely linked to cognitive abilities. Data from Hargrave et al. support the position of Blaskewitz et al. [3] [11].

3. Purpose statement

The purpose of this study was to determine whether MSVT performance in a sample of French Canadian schoolchildren differs from performance in other samples of schoolchildren presumed to have exerted good effort. A comparison was made between a sample of French Canadian schoolchildren who had undergone a psycho-educational evaluation

between January 2006 and December 2010 and the Normative samples comprising children from the manual, as well as with other groups of children available, including clinical samples of children. Another procedure, the Reliable Digit Span (RDS), was used with a modified cut-off for comparison and cross-validation of its use with children [1]. The comparison was made with variables such as performance on the effort subtests of the MSVT, scores on the RDS (when available), age, and reading levels.

4. Significance

By demonstrating that most children with learning difficulties and behavior problems can pass the MSVT, this study strengthens the validity of using the test with adults and children with normal cognitive abilities who are making a poor effort and producing invalid test results. According to Flaro, Green, and Allen, significant effects of suboptimal effort on test score validity have been reported in both adults and children with no known incentive to perform below their capacities. There are many reasons apart from money why some people might put forth inadequate effort during testing. Although most schoolchildren are not assessed in the context of litigation or when seeking special accommodations, performance may still be affected by external and internal incentives to perform below capacity (i.e., to put forth less than full effort) during a psycho-educational assessment, as well as task avoidance, or low self-efficacy.

Flaro et al. reported that a significant minority of children were found to produce such low scores on the effort measure Computerized Assessment of Response Bias (CARB), and the Word Memory Test (WMT) that “their ability test results would have to be regarded as being of doubtful validity” [6].

Kirkwood and Kirk examined the performance of 193 pediatric patients, aged 8 through 17 who had sustained mild traumatic brain injury [14]. The base rate of failure was 17% (33 children out of 193, with only one child whose lack of effort could be attributed to litigation), suggesting that a sizable minority of children are capable of putting forth suboptimal effort during a neuropsychological exam, even in the absence of apparent incentives [14].

Kirkwood, Yeates, Randolph, and Kirk studied a larger sample of 276 children referred to an outpatient concussion program at a children’s hospital in the United States [13]. The group was comprised of the same subgroup from the previously described study. Of the 276 participants, 51 (18.5% of the sample) failed at least one of the three primary validity indices of the MSVT. None of the children’s parents were seeking disability compensations or had external incentives to appear impaired. No background or injury-related variable differentiated

those who passed the MSVT from those who failed. Performance on the MSVT was correlated significantly with performance on all ability-based tests and explained 38% of the total ability-based test variance. Participants failing the MSVT performed significantly worse on nearly all neuropsychological tests, with large effect sizes apparent across most tests [13]. Kirkwood et al. concluded that the results provided compelling evidence that practitioners should add objective Symptom Validity Tests to the evaluation of school-aged youth, even when secondary gain issues are not readily apparent [13].

Children who think they lack the ability to perform, such as children with learning disabilities, may not produce optimal effort [2]. Moreover, lying and deception, relating to an overarching negative attitude towards authority, are quite commonly associated with certain diagnoses, such as oppositional defiant disorder (ODD) and conduct disorder (CD). Children with learning problems and/or ODD/CD abound in the caseload of a typical school psychologist. Objective effort-testing tools would certainly be a worthwhile addition to the traditional psycho-educational assessments tests since they would help validate the results.

This study added one comparison group—a clinical sample of 61 French Canadian children who had undergone a psycho-educational assessment between 2006 and 2010—to the existing MSVT and RDS data available on children. Digit Span, from which the Reliable Digit Span procedure is derived, is usually administered in all Wechsler Intelligence Scales tests, a gold standard in psycho-educational assessments. This research will compare the MSVT and RDS scores from a sample of children with the scores from other samples to help substantiate the use of these measures with neuropsychological evaluations. The study will also provide information for considering the use of these measures in psycho-educational assessment of children.

There has been little published research looking at measures of effort with children and adolescents. Yet neuropsychological researchers view this line of research—how children and adolescents perform compared with adults—as important [4]. Apart from the data available in the MSVT manual and appendices, seven studies included testing data related to children's effort after the manual was published: Teichner, Waid, and Buddin, Chafetz, Abrahams, and Kholmaier, Gill, Green, and Flaro, Carone, Blaskewitz, Merten, and Kathmann, Kirkwood et al. and Kirkwood and Kirk [3] [4] [5] [9] [13] [14] [20].

These studies reported pass/fail rates of children comparable to those of adults making good effort regardless of age or abilities, with the exception of Social Security claimants, who had a much higher failure rate [5]. Kirkwood and Kirk, who administered the MSVT to pediatric patients with

mild brain injury, did not find the pre-concussion diagnosis of a reading disorder contributed to the failure rate [14]. There is, however, no mention of reading levels for the children who successfully passed the effort measures (in the MSVT manual or subsequent studies) or for the minority of children who failed them. Knowing the child's reading-level is an important element addressed by this current research. In addition, the groups of children presenting specific impairments described in the manual were usually small-*N*, ranging in size from 9 to 46 children.

Symptom validity tests are not yet widely used in children's psycho-educational assessments, although they have been reported to be useful in psycho-educational assessments of college students [7] [18]. Sullivan, May, and Galbally investigated the failure rate in WMT effort measures of 66 self-referred college students who had undergone assessment for ADHD, a learning disability (LD), or both [18]. The authors reported that 22.4% of the clients who self-referred for an ADHD assessment, 15.4% of the clients who self-referred for an LD assessment, and 9.4% of the clients who self-referred for both ADHD and LD assessments failed at least one effort measure. These figures are relatively high, considering the failure rate of individuals making full effort (or at least claiming to) on the WMT is less than 2%.

Harrison, Larochette, and Edwards administered the WMT to a group of learning-disabled children undergoing psycho-educational assessments [12]. Overall, 89% of the students passed and 11% failed the effort measures. The authors found that WMT measures were significantly positively correlated to the WIAT word-reading score and concluded that, although the WMT seems to be an appropriate measure of effort for the majority of reading-disabled children, it is not an appropriate measure of effort in children with profound and well-documented word decoding impairments. The sample used by Harrison, Larochette, and Edwards is closest to the sample used in this study [12]. Since the MSVT is a shortened version of WMT, the findings of this study were expected to be comparable.

5. Methodology

A comparison was made between effort measures of the MSVT (IR, DR and CNS scores) of a convenience sample of 61 French Canadian schoolchildren who had undergone a psycho-educational evaluation between January 2006 and December 2010 and the Normative samples comprising children from the manual, as well as with other groups of children available. Another procedure, the Reliable Digit Span (RDS), was used with a modified cut-off for comparison and cross-validation of its use with children [1].

Of the 61 children from Alberta and the Northwest Territories ranging from grade 1 to 12 and from age 6 to 18, 51 children were Caucasian, had either French or English as dominant language but spoke both languages. 25 children from the sample were less than 11 years of age. 36 children were 11 years of age or above. 49 children read at grade 3 level or above. 12 children read below grade 3 level (as measured by their score on WIAT-II Word Reading at the time of the assessment). 25 were under 11 years of age and 36 were 11 years of age or older.

Of the 61 children, 2 received a Tourette Syndrome diagnosis; 17 had a primary diagnosis of ADHD and LD; 4 had a primary diagnosis of ADHD, ODD, and LD; 3 had a diagnosis of ADHD, ODD, and CD (no LD); 4 had LD and a suspected diagnosis of PDD; 10 had LD only; 9 were referred for gifted assessment; 7 were slow learners (borderline IQ or mild Mental Retardation); 2 had fine motor skills problems and slow processing speed (no other diagnoses); 3 remained without diagnosis.

The French Canadian sample was compared with the previously described research samples reported in the MSVT manual and one from a research paper published subsequently by Teichner, Waid, and Buddin [20]. The data available on the children in the MSVT test manual are as follows:

1. Preliminary Normative MSVT sample of 55 healthy Canadian children tested in English (grade 2, N = 20, grade 3, N = 16, grade 5, N = 19). The mean IR, DR, and CNS were 99.4%, 99.2%, and 98.6%, respectively.
2. Twenty-six learning-disabled children, aged 12, mean Full Scale IQ = 91. The mean IR, DR, and CNS were 100%.
3. Forty-six children with Conduct Disorder, aged 14, mean Full Scale IQ = 92. The mean IR, DR, and CNS were 100%, 99%, and 99%, respectively.
4. Thirty-three children with ADHD, aged 11, mean Full Scale IQ = 92. The mean IR, DR, and CNS were 98%, 96%, and 96%, respectively.
5. One hundred and two American children with various clinical conditions [20]. They obtained a 98.6% mean score on IR and 95.8% on DR. No CNS score was reported.
6. Thirty-six Brazilian good-effort volunteers (aged 6 to 10) who were administered the MSVT in Portuguese. The mean IR, DR, and CNS were 94.6%, 99.3%, and 93.9%, respectively.
7. Thirty-four Brazilian good-effort volunteers (aged 11 to 15) who were administered the MSVT in Portuguese. The mean IR, DR, and CNS were 96.3%, 99.6%, and 95.9%, respectively.
8. Thirty-five children with Fetal Alcohol Syndrome (FAS), aged 14, mean Full Scale IQ = FSIQ 85. The mean IR, DR, and CNS were 100%, 99%, and 98%, respectively.

Except for the Normative sample, the groups of children that were used for comparison from the manual, as well as the group of American children with various conditions (Teicher, Waid & Buddin, 2005) were convenience samples. No other data were reported for these groups.

There were 6 research questions.

1: What was the difference between the mean Immediate Recognition (IR) scores for a French Canadian children's sample compared to other comparative groups?

2: What was the difference between the mean Delayed Recognition (DR) scores for a French Canadian children's sample compared to other comparative groups?

3: What was the difference between the mean Consistency (CNS) scores for a French Canadian children's sample compared to other comparative groups?

4: What was the difference between the two levels of age groups (less than 11 years of age, 11 years and older) of the French Canadian sample on the three MSVT mean effort subtests scores (IR, DR, and CNS)?

5: What was the difference between the two levels of reading (WIAT-II Word Reading grade equivalent)?

6: What was the relationship between the MSVT effort measures and the Wechsler DS score (RDS) in the French Canadian sample?

To test the hypothesis for questions 1, 2 and 3, the one-sample t-test was used to see if two samples means were different from each other. The two samples that the means are based on were taken from different individuals who have not been matched. There is no non-parametric equivalent for the t-test: despite the fact that the distributions were moderately negatively skewed, we decided to use the one-sample t-test since all were skewed in the same direction and attempts to transform the data, as suggested by Tabachnik and Fidell, resulted in inverted distributions [19]. IVs were group memberships, DVs were on an interval scale (IR, DR and CNS MSVT effort measures scores). Statistical analyses were run twice: once with the original sample of 61 children, and once with the 49 children who were able to read at grade 3 level or above.

For questions 4 and 5, a MANOVA was planned but could not be used, due to violations of assumption of normality. By transforming the DVs into ordinal scale of measurement, a comparison was made between mean ranks and Mann-Whitney U was used instead.

Point biserial correlation r_{pb} was used for question 6. Level of significance for all statistics was set at .05.

The demographic information of participants, such as age and grade, were recorded in the Statistical Program for Social Sciences (SPSS). Scores on MSVT Immediate Recognition (IR), Delayed Recognition (DR), and Consistency (CNS); raw Digit Span on the Wechsler tests; WIAT-II Word Reading subtest scores; and diagnoses were entered. Descriptive statistics including (a) frequency counts, (b) percentages, and (c) means were reported for demographic variables in order to compare them with other samples. Available data for comparative samples (MSVT IR, DR, and CNS; Digit Span raw and standardized scores; Reliable Digit Span; as well as age and diagnoses) were also entered.

6. Results

The mean MSVT effort measure results for the 61 French Canadian participants were as follows: Immediate Recall (IR) 95.25, for Delayed Recall (DR) 96.23, and Consistency (CNS) 94.43. Of the 61 participants, 86.9% of the participants passed the MSVT (i.e., obtained a score above 85% on IR, DR and CNS as defined by the Manual), while 13.1% (8 out of 61) failed these same measures. On the Reliable Digit Span (RDS) effort measure, 63.9% passed and 36.1% (22 out of 61) failed. When the children who read below grade 3 were removed (N = 49) for a second analysis, the means for IR, DR and CNS were 97.65, 96.67 and 97.14, respectively. The failure rate on the MSVT dropped to 6.1%. On the Reliable Digit Span (RDS) effort measure, of the 61 children comprised in the sample, 63.9% passed and 36.1% failed.

Question 1 referred to the difference between the mean Immediate Recognition (IR) scores for a French Canadian children’s sample compared to other comparative groups of children (found in the MSVT manual and other publications).

Table 1. One-sample t-test Combined Results for Immediate Recall by Group/Sample Population (Comparative) Mean Sample Mean n = 61

Group/Sample	Mean ^a	Mean ^b	t	p
Normative	98.80	95.25	2.76	.008
LD	100.00	95.25	3.69	<.001
CD	100.00	95.25	3.69	<.001
ADHD	98.00	95.25	2.14	.037
American Children	98.60	95.25	2.60	.012
Brazilian, age 6-10	94.60	95.25	.50	.618
Brazilian, age 11-15	96.30	95.25	.82	.416
FASD	100.00	95.25	3.69	<.001

The French Canadian sample had significantly lower mean IR scores than the other comparison groups, except for the two Brazilian children groups.

When French Canadian children reading below grade 3 were removed, the sample was not significantly different from the Normative, ADHD, and American children groups. French Canadian children reading at or above grade 3 had mean IR significantly higher than two Brazilian children groups. French Canadian children reading at or above grade 3 had mean IR significantly lower than LD, CD, and FASD groups (group means = 100).

Question 2 referred to the difference between the mean Delayed Recognition (DR) scores for a French Canadian childrens sample compared to other comparative groups of children.

Table 2, children reading below grade 3 removed One-sample t-test Combined Results for Immediate Recall by Group/Sample Population (Comparative) Mean Sample Mean of Children who read at or above grade 3 level, n = 49

Group/Sample	Mean ^a	Mean ^b	t	p
Normative	98.80	97.65	1.64	.108
LD	100.00	97.65	3.38	.002
CD	100.00	97.65	3.35	.002
ADHD	98.00	97.65	.50	.623
American Children	98.60	97.65	1.35	.183
Brazilian, age 6-10	94.60	97.65	3.79	<.001
Brazilian, age 11-15	96.30	97.65	2.36	.022
FASD	100.00	97.65	3.35	.002

The French Canadian sample had significantly lower mean DR scores than children with learning disabilities samples and both Brazilian children samples, but not from the Normative, CD, FASD, ADHD samples, nor from the American children samples. When the children reading below grade 3 level were removed, the mean DR score was not significantly different than the two Brazilian and from the Normative samples. The French Canadian mean DR score was higher than ADHD and American children samples but was significantly lower than the LD group.

Question 3 referred to the difference between the mean Consistency (CNS) scores for a French Canadian childrens sample compared to other comparative groups of children.

Table 3. One-sample t-test Combined Results for Delayed Recall by Group/Sample Population (Comparative) Mean Sample Mean n = 61

Group/Sample	Mean ^a	Mean ^b	<i>t</i>	<i>p</i>
Normative	98.50	96.23	1.63	.108
LD	100.00	96.23	2.71	< .001
CD	99.00	96.23	1.99	.051
ADHD	96.00	96.23	0.17	.230
American Children	95.80	96.23	0.31	.759
Brazilian, age 6-10	99.30	96.23	2.21	.031
Brazilian, age 11-15	99.60	96.23	2.42	.018
FASD	99.00	96.23	1.99	.051

Question 4 referred to the difference between the two levels of age groups (less than 11 years of age, 11 years and older) of the French Canadian sample on the three MSVT mean effort subtests scores (IR, DR, and CNS).

Table 4. children who read below grade 3 removed One-sample t-test Combined Results for Delayed Recall by Group/Sample Population (Comparative) Mean Sample Mean n = 49

Group/Sample	Mean ^a	Mean ^b	<i>t</i>	<i>p</i>
Normative	98.50	98.67	0.36	.719
LD	100.00	98.67	2.77	.008
CD	99.00	98.67	0.68	.498
ADHD	96.00	98.67	5.59	<.001
American Children	95.80	98.67	6.00	<.001
Brazilian, age 6-10	99.30	98.67	1.31	.197
Brazilian, age 11-15	99.60	98.67	1.84	.059
FASD	99.00	98.67	0.68	.498

There were significant differences on IR, DR, and CNS mean ranks between the less than 11 years of age and the 11 years and older groups. When the children who read below grade 3 were removed, there were no longer significant differences on IR, DR mean ranks between the less than 11 years of age and the 11 years and older groups.

Question 5 referred to the difference between the two levels of reading (WIAT-II Word Reading grade equivalent below 3.0, WIAT –II Word Reading grade equivalent at or above 3.0) of the French Canadian sample on the three MSVT mean effort subtests scores (IR, DR, and CNS).

Table 5. One-sample t-test Combined Results for Consistency by Group/Sample Population (Comparative) Mean Sample Mean n = 61

Group/Sample	Mean ^a	Mean ^b	<i>t</i>	<i>p</i>
Normative	97.70	94.43	2.10	.040
LD	100.00	94.43	3.57	< .001
CD	99.00	94.43	2.93	.005
ADHD	96.00	94.43	1.01	.317
American Children	N/A	N/A	N/A	N/A
Brazilian, age 6-10	93.90	94.43	.34	.737
Brazilian, age 11-15	95.90	94.43	.95	.349
FASD	99.00	94.43	2.93	.005

There were significant differences on IR, DR, and CNS mean ranks between the Word Reading grade equivalent below 3.0 group and the Word Reading grade equivalent 3.0 or above group.

Question 6 referred to the strength of the relationship between the failure of MSVT effort measures and the failure of RDS score in the French Canadian sample [1]. A point-biserial correlation was calculated to answer research question 6. The point-biserial correlation is a special case of the Pearson *r* correlation coefficient that is applied to dichotomous variables, in this case, pass or fail on effort measures (either IR, DR, or CNS) and pass or fail on Reliable Digit Span. The cut-off score for a pass on RDS was above 6. A score of 6 or below was a fail.

For the sample with *N* = 61, there was a moderately positive correlation (*r_{pb}* = .315, *p* < .05) between passing/failing MSVT effort measures and passing/failing Reliable Digit Span, indicating that passing or failing MSVT effort measures is moderately related to passing or failing Reliable Digit Span. When the 12 participants who read below grade 3 level were removed, the correlation between pass/fail MSVT and pass/fail Reliable Digit Span became insignificant.

7. Discussion

Limitations for this research include the fact that this is a convenience clinical sample of French Canadian children and did not make use of a control group or other comparison group for the population of French Canadian children. Comparisons between mean effort scores of the different samples, as seen in the previous sections, must be interpreted with caution, considering the normality assumption violation. Also, we do not know the reading level of the 8 samples of children used for comparison.

Differences were found between the means of the French Canadian sample and several other children

samples (clinical, normal and Normative) although fewer differences between the means of the French Canadian children and the other eight MSVT samples were found when 12 participants reading below grade three were removed from the sample and the statistical tests performed a second time with the remaining sample of 49 children. Once the reading level factor was controlled by removing participants with an inadequate reading level from the sample in the second data analysis, the 6.1 % failure rate for the French Canadian sample was close to that of presumably good-effort samples, including the Normative sample, as well as most other clinical samples. In fact, there was no significant difference between the pass/fail rate of the effort measures between the French Canadian and the Normative sample (5% failure rate). The vast majority of students passed the MSVT effort measures, even the participants exhibiting behavioral problems, as was the case for the other children's clinical samples and in the literature documenting the use of MSVT with children with and without impairments [3] [4] [8] [9].

Green (2012, personal communication) suggested considering evidence of a Genuine Memory Impairment Profile (GMIP) in failing protocols (i.e., when the average of IR, DR, and CNS exceeds the average of PA and FR by at least 20 points). In this study, all but one child who failed MSVT demonstrated a GMIP profile. These results were consistent with the literature documenting the use of MSVT and WMT with adults, with and without impairments, and supported the assertion that insufficient effort was likely the only explanation for failure on MSVT when the participants had the required reading level [8].

There may be an impact on the scores when MSVT is administered in another language than English, as the French Canadian and the Brazilian samples both demonstrate significantly different performance means relative to the Normative sample. A sizable proportion of the French Canadian children were administered the MSVT in French. This was not anticipated, but future research might consider whether differences are attained when the MSVT is administered in various languages.

One of the most recent studies of effort testing with children, using WMT, found that 40% of children who read below grade three level failed, while 60% of students passed [12]. In the French Canadian sample, 5 out of 12 students (41%) reading below grade three level failed the MSVT effort measures, while 7 out of 12 (59%) students reading at or below grade three level passed, yielding a very similar picture. An examination of each case suggested that the participant with the poorest reading levels tended to be the youngest, and indeed, there was a moderately significant positive

correlation ($r_{pb} = .426, p < .01$) between age and reading level, as was expected.

There are clinical choices to be considered when using MSVT with clients who read below grade 3 in the course of an assessment. All but one of the French Canadian children who read below grade 3 level (the level Green has determined necessary to be able to pass MSVT) presented with a GMIP profile [8]. Green indicated in the manual that accommodations could be made with these children (such as reading the words out loud on the screen with the child during administration) [8]. A nonverbal test – the NV-MSVT – can also be used.

Failure rate on the RDS (36.1%, with a cut-off of 6) was moderately but significantly correlated with failure on effort measures with the full sample of 61 French Canadian children. It would have been interesting to measure the relationship between LD diagnosis and a pass or fail on RDS, as poor reading has been associated with poorer working memory on the WISC-IV [17]. RDS would not, then, be considered a good measure of effort for this population.

LD was present either as a main diagnosis or as co-morbidity for a substantial number of children (who would not be a pure sub-sample of children with a learning disability). Further research assessing the mean raw Digit Span score of learning-disabled children would be useful to determine if RDS should be used with this population, and what cut-off would be recommended.

8. References

- [1] Babikian, T., K. Bauer Boone, P. Lu, and G. Arnold, "Sensitivity and specificity of various Digit Span scores in the detection of suspect effort", *The Clinical Neuropsychologist*, 2006, issue no. 20, pp. 145-159.
- [2] Baird, G. L., W.D. Scott, E. Dearing, and S.K. Hamill, "Cognitive self-regulation in youth with and without learning disabilities: academic self-efficacy, theories of intelligence, learning vs. performance goal references, and effort attributions", *Journal of Social and Clinical Psychology*, 2009, issue no. 28(7), pp. 881-908.
- [3] Blaskewitz, N., T. Merten, and N. Kathman, "Performance of children on symptom validity tests: TOMM, MSVT, and FIT", *Archives of Clinical Neuropsychology*, 2008, issue no. 23, pp. 379-391.
- [4] Carone, D., "Children with moderate/severe brain damage/dysfunction outperform adults with mild-to-no brain damage on the Medical Symptom Validity Test", *Brain Injury*, 2008, pp. 960-971.
- [5] Chafetz, M.D., J.P. Abrahams, and J. Kohlmaier, "Malingering on the social security disability consultative exam: A new rating scale", *Archives of Clinical Neuropsychology*, 2007, issue no. 22, pp. 1-14.

[6] Flaro, L., P. Green, and N. Blaskewitz, "Die Bedeutung der Beschwerdvalidierung im Kindesalter" [The importance of symptom validity testing in children: WMT & MSVT], *Praxis der Rechtspsychologie*, Germany, 2007, issue no. 17(1), pp. 125-139.

[7] Frazier, T.W., A.R. Frazier, R.M. Busch, M.A. Kerwood, and H.A. Demaree, "Detection of simulated ADHD and reading disorder using symptom validity measures", *Archives of Clinical Neuropsychology*, 2008, issue no. 23, pp. 501-509.

[8] Green, P., *Green's Medical Symptom Validity Test (MSVT) for Microsoft Windows, User's manual*, Green's Publishing, Edmonton, Canada, 2004.

[9] Green, P., and L. Flaro, "Word Memory Test performances in children", *Child Neuropsychology*, 2003, issue no. 9(3), pp. 189-207.

[10] Greiffenstein, M.F., J.W. Baker, and G. Thomas, "Validation of malingered amnesia measures with a large clinical sample", *Psychological Assessment*, 1994, issue no. 6(3), pp. 218-224.

[11] Hargrave, D.D., M.W. Kirkwood, and J.W. Kirk, "Sensitivity and specificity of digit span scores in identifying suboptimal effort after pediatric mild traumatic brain injury" [Abstract], *The Clinical Neuropsychologist*, 2010, issue no. 24(4), p. 625.

[12] Harrison, A., A.C. Larochette, and M. Edwards, "Word Memory Test performance in children with reading disabilities", poster session presented and the NAN annual conference, New Orleans, LA., November, 2009.

[13] Kirkwood, M.W., K.O. Yeates, C. Randolph, and J.W. Kirk (2011) "The implication of Symptom Validity Test failure for ability-based test performance in a pediatric sample"; *Psychological Assessment; Advance online publication*. doi: 10.1037/a0024628.

[14] Kirkwood, M.W., and J.W. Kirk, "The base rate of suboptimal effort in a pediatric mild TBI sample: performance on the Medical Symptom Validity Test", *The Clinical Neuropsychologist, iFirst*, 2010, pp. 1-13.

[15] Richman, J., P. Green, R. Gervais, L. Flaro, T. Merten, R. Brockhaus, and D. Ranks, "Objective tests of symptom exaggeration in independent medical examinations", *Journal of Occupational and Environmental Medicine*, 2006, issue no. 48, pp. 303-311.

[16] Ruff, R.M., T. Wylie, and W. Tennant, "Malingering and malingering-like aspects of mild closed head injury", *Journal of Head Trauma Rehabilitation*, 1993, issue no. 8, pp. 60-73.

[17] Rury Smith, D. (2004) "Making the reading connection", http://www.pearsonassessments.com/NR/rdonlyres/BD9F8859-88A4-4F39-B39E-AB3A92108908/0/WISCIV_and_Reading_Handout.pdf

[18] Sullivan, B.K., K. May, and L. Galbally, "Symptom exaggeration by college adults in Attention-Deficit

Hyperactivity Disorder and Learning Disorder assessments", *Applied Neuropsychology*, 2007, issue no. 14(3), pp. 189-207.

[19] Tabachnik, B.G., and L.S. Fidell, *Using multivariate statistics* (4th ed.), Allyn and Bacon, 2001.

[20] Teichner, G., R. Waid, and W.H. Buddin, *The Medical Symptom Validity Test (MSVT): Data from a sample demonstrating a range of neurological and psychiatric disorders*, poster session presented at the NAN Annual Conference, Tampa Bay, FL, October, 2005.

9. Acknowledgements

This research paper was first submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Educational Psychology, School Psychology Emphasis, at Northern Arizona University, May 2012.

I would like to thank the members of my dissertation committee for having "stuck" with me during this very long and difficult process. I would like to thank especially Dr. Bohan, chair of the committee, who believed in me to the end, as well as Dr. Gervais, from Neurobehavioral Associates in Edmonton, Alberta, who introduced me to SVTs, instructed me on their use, and finally accepted to become a member of the committee.

I am especially grateful to Dr. Green, from Edmonton's Neurobehavioral Associates, who graciously made the MSVT available to me and directed me towards relevant resources. In addition, I thank Dr. Horn for giving me very good advice about the use of statistics. I would also like to thank the Conseil scolaire Centre-Nord and the Réseau provincial d'adaptation scolaire for their permission to use the data for this study and for facilitating access to that data.