Differential diagnosis in the memory clinic: Exploring the value of improved neuropsychological examination
Rienstra, A.

Citation for published version (APA):
CHAPTER 3

Validation of symptom validity tests using a 'child-model' of adult cognitive impairments

Anne Rienstra, Pauline E.J. Spaan and Ben Schmand

Archives of Clinical Neuropsychology, 2010; 25: 371-382
ABSTRACT
Validation studies of symptom validity tests (SVTs) in children are uncommon. However, since children’s cognitive abilities are not yet fully developed, their performance may provide additional support for the validity of these measures in adult populations.

Four SVTs, the Test of Memory Malingering (TOMM), the Word Memory Test (WMT), the Amsterdam Short Term Memory (ASTM) test and the Word Completion Memory Test (WCMT), along with several neuropsychological instruments were administered to 48 Dutch school children aged 7 to 12. All children scored above the established adult cut-offs on the TOMM and the WMT. They could pass the ASTM test if their reading skills were at a level equivalent to that of 9 year olds. All children passed our criterion of a negative WCMT score. However, the WCMT does seem sensitive to level of verbal fluency. Implications for the applicability of these SVTs in adult populations are discussed.
1. INTRODUCTION
Clinicians and researchers have become increasingly concerned about the impact of suboptimal effort on neuropsychological test performance, because this may reduce the validity of test results (Constantinou et al., 2005; Green & Grant, 2001; Slick et al., 2004). Consequently, it is recommended that objective tests of effort be used, especially when compensation claims are involved (Bush et al., 2005; Iverson & Binder, 2000). In recent years, multiple methods and measures have been developed to assess suboptimal performance (Bianchini et al., 2001; Larrabee, 2003). When using effort measures in clinical practice, clinicians should ensure that their validity has been adequately investigated. There is, however, no gold standard to determine with certainty whether examinees are exerting suboptimal performance other than the effort tests being investigated (Cercy et al., 1997). The most convincing results concerning the sensitivity of a SVT would be obtained from examinees who admit to exert insufficient effort. Unfortunately, those people will rarely make such confessions. Therefore, it is impossible to directly determine the criterion validity of SVTs.

Until now, the validation of effort measures has generally been done with four different types of studies. First, normative research has been conducted with cognitively intact adults (e.g., Bolan et al., 2002; Iverson et al., 1999; Rees et al., 1998; Schagen et al., 1997). Second, the specificity of SVTs has been studied in clinical populations consisting of patients with neurological disorders (e.g., Howe et al., 2007; Macciocchi et al., 2006; Mathias et al., 2002) and psychiatric syndromes (e.g., Ashendorf et al., 2004; Gorissen et al., 2005; Weinborn et al., 2003). Third, in the absence of a gold standard for malingering (i.e., confessions of malingerers), numerous analogue studies with experimental malingerers have been conducted (e.g., Brockhaus & Merten, 2004; Haines & Norris, 2001; Iverson & Franzén, 1996; Merten et al., 2005; Rees et al., 1998) to establish the sensitivity of SVTs. For the same reason, much research has been conducted in litigating populations (e.g., Gervais et al., 2004; Green et al., 1999; Stevens et al., 2008).

Whereas most SVTs have been validated by testing adults from various populations, validation studies using children are scarce. One may assume, however, that children's performance on SVTs provides additional support for the specificity of these tests. Most cognitive abilities increase as a function of age (e.g., Anglin et al., 1993; Fry & Hale, 1996; Rueda et al., 2004; Romine & Reynolds, 2005). Therefore, the performance of children on cognitive tests, including SVTs, can serve as a 'child model' of adult cognitive impairments, in a similar way as animal models are being used to investigate human diseases. If it can be shown that the immature cognitive abilities of
young children do not interfere with performance on SVTs, then this would provide
evidence in support of the assumption that effort measures are relatively insensitive to
mild cognitive disorders at an adult age.

In earlier reports it has been described that children may not always put forth
maximal effort during evaluations (Baron et al., 1995; Lu & Boone, 2002). Furthermore,
it has been shown that they can be coached successfully to feign or exaggerate
neuropsychological impairment and that invalid data resulting from deception are
not easily detected by clinicians (Faust et al., 1988; Faust et al., 1988; Hall & Pritchard,
1996). As a result of these works, previous studies on SVT performance in children
have focused primarily on the question whether well-established SVTs could be use-
ful in the evaluation of effort in paediatric assessments. Constantinou and McCaffrey
(2003) reported that the performance of children as young as 5 years of age on Trial
2 of the Test of Memory Malingering (TOMM; Tombaugh, 1996) was comparable to
that of healthy adults. Subsequent studies in the USA and Germany (Blaskewitz et
al., 2008; Nagle et al., 2006) replicated this finding in children aged 6 years. Finally,
Donders (2005) demonstrated that even in a sample of children with a wide range
of clinical diagnoses, more than 90% of the 6 to 8 year olds exceeded the established
adult cut-offs for sufficient effort. The same pass rate was found in children between 6
and 17 years of age with a diagnosis of epilepsy (MacAllister et al., 2009).

Also for the Word Memory Test (WMT; Green et al., 1996), several researchers
have investigated whether it is potentially useful as a measure of effort in children.
Based on data from 135 children between the ages of 7 and 18 years, again with a vari-
ety of clinical diagnoses, Green and Flaro (2003) suggested that most children aged 7
years or above with at least a grade 3 reading level can pass the WMT using the adult
criteria. These results supported the authors’ previous findings that 90% of children
aged 8 to 18 years could pass all WMT effort subtests (Flaro et al., 2000). Courtney
et al. (2003), however, found that performance on the WMT effort measures was sig-
nificantly related to age and reading level, with only children of 11 years and older
producing results similar to those of adult participants.

The present investigation aimed to determine whether the validity of several effort
measures that are frequently used in clinical practice could be bolstered with the per-
formance of children. The questions of interest were (a) whether young children are
able to pass the adult criteria of SVTs, (b) whether their performance is comparable to
that of healthy adults, and (c) which cognitive immaturities might be responsible in
case children would not be capable of successfully completing the SVTs.

Dutch schoolchildren between the ages of 7 and 12 years were administered four SVTs
Exploring the value of improved neuropsychological examination along with several neuropsychological instruments. The WMT is one of the most widely researched and most sensitive effort tests used in adults (Hartman, 2002). The TOMM is a non-verbal SVT with high specificity and little impact of cultural factors. Furthermore, the TOMM has been described as the effort measure most commonly used by neuropsychologists who are expert in handling injury cases (Slick et al., 2004). Two other SVTs that may add valuable information are the Amsterdam Short-Term Memory test (ASTM test; Schmand & Lindeboom, 2004) and the Word Completion Memory Test (WCMT; Hilsabeck & LeCompte, 1997). The ASTM test is the first SVT that masks the obvious a priori probability of a correct response (which is 50% for most SVTs) by requiring the examinees to select three out of five target words on each trial (Lezak et al., 2004). The WCMT is a priming/implicit memory measure, which makes it the first SVT that has the potential to show that participants are deliberately exerting suboptimal effort. This task is based on a process dissociation procedure developed by Jacoby (1991). Process dissociation provides a framework for measuring the independent contributions of consciously controlled and automatic memory processes. The rationale behind the approach is that “conscious control can be measured as the difference between performance when a person is trying to as compared with trying not to use information from some particular source” (Jacoby, 1991; p.527). On the WCMT, participants are first instructed to complete word stems with words from a previously studied list. On a second task, examinees are asked to complete word stems with words that were not included in a previously studied list. The difference score between these conditions is hypothesized to be an estimate of the conscious control of participants over their memory for the previously studied words.

In the light of previous research (Blaskewitz et al., 2008; Constantinou & McCaffrey, 2003; Donders, 2005; Green & Flaro, 2003; Nagle et al., 2006), it was hypothesized that all children, irrespective of their age, would be able to pass the TOMM and the WMT according to the established adult cut-offs, and that their performance would be comparable to that of healthy adults. In the absence of relevant previous studies, no predictions were made for the ASTM test and the WCMT.

If it would turn out that children fail these SVTs, it was anticipated that performance on the WMT and the ASTM test would be affected by low reading level. This anticipation was based on the findings of earlier studies that performance on the WMT was significantly related to reading ability (Courtney et al., 2003; Green & Flaro, 2003). Since the presentation time of the ASTM stimuli is limited, the same was expected for this task. For the ASTM test, as well as for the WCMT, it was further hypothesized that test scores would be related to level of vocabulary, because it is easier to remember
familiar than unfamiliar words (Brandt et al., 2005; Hulme et al., 1991). In addition, it was expected that WCMT performance would be associated with verbal fluency because the WCMT requires the generation of words to complete word stems. Finally, consistent with previous findings (Merten et al., 2007) of low correlations of SVTs with the Trail Making Test (TMT; Reitan, 1992) and Digit Span (Wechsler, 1997), we hypothesized that attention and concentration would not affect SVT performance.

2. METHODS
2.1. PARTICIPANTS
In this study 48 Dutch children between the ages of 7 and 12 years ($M = 119.3$ months, $SD = 19.3$) participated. Children were recruited in grades 2 through 6 from two different elementary schools. On both schools, four children (two boys and two girls) were randomly selected for each age. Exclusion criteria were: (a) learning disabilities based on information acquired from the child's school record; (b) a history of physical (including neurological) conditions or psychiatric disorders that could interfere with cognitive performance; (c) Dutch as a second language.

2.2. PROCEDURE
Permission for the conduction of the study was obtained from the Institutional Review Board of the Psychology Research Institute of the University of Amsterdam. A letter with a detailed description of the study was sent to the parents or other legal guardians of the children who were randomly selected to participate. Written informed consent was given by the parents. Participation was strictly voluntary. At their schools, the children were administered a test battery containing several neuropsychological instruments (see Materials section). Data regarding reading ability were collected by gathering the most recent scores on a speeded reading test (see Materials) from the children's school records. Children were tested individually by the first author. The complete assessment took approximately an hour and a half. The children were instructed to perform optimally on all measures. Independent of their performance, all participants were allowed to choose a small present at the end of the test session.

The order of tests was fixed to the sequence described below (with the WMT DR subtest and the TOMM Retention Trial administered subsequent to the TMT and the WCMT, respectively). The sequence was chosen in such a way that potential interference of similar test modality was prevented, and that tests measuring effort were alternated with those assessing cognitive abilities. However, to control for effects of...
test order and fatigue, the initial and final tests were counterbalanced across participants. Half of the children started the test session with the ASTM test, whereas the other half started with the WCMT. Furthermore, two different versions of the WCMT were used. To investigate the possible influence of their composition, the two WCMT word lists were counterbalanced across the Inclusion and Exclusion subtests. Finally, to compare the equivalence of the computerized and orally administered WMT, the oral form of the WMT was given to one half of the children and the computerized test was given to the other half. On both schools and for each age level, at random two out of four children were given the oral form of the WMT.

2.3. MATERIALS

Three-Minutes-Test (3-MT; Verhoeven, 1995). This speeded reading test is periodically administered by Dutch schools as part of a pupil monitoring system for 4- to 12-year-olds. The 3-MT consists of three reading cards of increasing difficulty with words of varying complexity. Children are instructed to read aloud as many words as possible in one minute. Reading level is determined by the total number of words that is read correctly. For this study, we only used the test score on card 3.

Amsterdam Short-Term Memory test (ASTM; Schmand & Lindeboom, 2004). The ASTM test is a 30-trial word recognition test. Each trial consists of five words from a particular semantic category (e.g., five vegetables), which participants are instructed to read aloud and to memorize. Subsequently, the participant has to solve a relatively simple addition or subtraction problem. Finally, participants have to select three out of five words (containing two semantically related foils) which they recognize from the first list.

Controlled Oral Word Association Test (COWAT; Benton & Hamsher, 1976). Participants are instructed to orally generate words that begin with the letters M and K in two separate, 1-minute trials, excluding proper nouns, numbers and the same words with a different suffix.

Word Memory Test (WMT; Green et al., 1996). The WMT is a word list learning task. In the computerized version, a list of 20 semantically related simple word pairs is presented twice. Subsequently, participants are required to select the 40 original words in a two-alternative forced-choice recognition test containing semantically related foils (the Immediate Recognition (IR) subtest). Without advance warning after a 30-minute delay, the Delayed Recognition (DR) subtest is administered in the same way as the IR subtest, except for using different foils. Based on the performance on the IR and DR subtests, the computer calculates the consistency score (CNS) as a percent
agreement in responses between these two trials. Any score on the IR, DR and CNS measures below the recommended cut-off point is indicative of suboptimal effort. The oral version consists of identical stimulus words and subtest procedures. The difference with the computerized WMT is that the words are read aloud to the participants in the learning phase and that they orally choose words from the words pairs that are read aloud in the recognition subtests.

Test of Memory Malingering (TOMM; Tombaugh, 1996). The TOMM is a two-alternative forced-choice recognition task consisting of three trials. In Trial 1, a total of 50 line drawings are presented for three seconds each. Subsequently, a recognition trial is administered, in which one of the previously presented pictures is shown again along with a new one. Participants are required to select the picture that was shown during the learning trial. Trial 2 is identical to Trial 1, except for containing different foil drawings. Approximately 15 minutes after Trial 2, the Retention Trial is administered, applying the same procedure except for not re-administering the target pictures. Again, different foil drawings are used.

Subtest Digit Span from the WISC-III (Wechsler, 1991). In the Digit Span subtest children are given sequences of numbers orally and are asked to repeat them, either as heard or in reverse order.

Trail Making Test (TMT; Reitan, 1992). The Trail Making test consists of two parts, A and B. In part A, examinees are asked to draw lines to connect consecutively numbered circles. Part B requires participants to connect consecutively numbered and lettered circles by alternating between the two sequences. Participants are urged to work as fast as possible. For this study, we used the Intermediate Version of the TMT which consists of 15 stimuli for both parts A and B.

Subtest Vocabulary from the WISC-III (Wechsler, 1991). In the Vocabulary subtest participants are asked to give definitions of words of increasing difficulty.

Word Completion Memory Test (WCMT; Hilsabeck & LeCompte, 1997). The WCMT has two subtasks, Inclusion and Exclusion. In the Inclusion subtest, each of 30 words is read aloud to the participants. Subsequently, the examinees repeat the word out loud and rate it for pleasantness. Then participants are asked to complete word stems with words from the previously presented list from memory (i.e., the inclusion task). In the Exclusion task, the participants repeat and rate 30 other words but are then instructed to complete word stems with words that were not included in the previously presented list (i.e., the exclusion task). This yields three scores: (1) an I score, which is the number of stems completed with words from the list presented in the Inclusion subtest; (2) an E score, which is the number of stems completed with
Exploring the value of improved neuropsychological examination

words from the list presented in the Exclusion subtest; and (3) an $R$ score, which is the difference between the $I$ and the $E$ score. This difference score is hypothesized to be a measure of the participants’ conscious control over their memory of the previously studied words (Hilsabeck et al., 2001). Persons who are simulating memory impairment are expected to obtain a negative $R$ score, because their intentionally poor memory performance would result in few word stems completed with words from the previously presented list on the Inclusion subtest and many stems completed with words from the list on the Exclusion subtest (Hilsabeck et al., 2001).

For this study, a Dutch translation of the original English WCMT was developed. To enable a comparison between the two language versions, we attempted to match the administration procedures as closely as possible. However, the Dutch WCMT is distinct from the English version in three aspects. First, contrary to the original Hilsabeck et al. (2001) study in which a cut-off of $<9$ for the $R$ score was used, an $R$ score equal to or lower than zero was defined as the cut-off. This is more consistent with the assumption that malingerers who are exerting suboptimal effort deliberately obtain a negative $R$ score. Second, whereas the English version uses a 7-point scale to rate the words for pleasantness, a scale ranging from 1 to 10 was applied in the Dutch WCMT. This was done to make the task of rating words for pleasantness more accessible to children. Giving marks ranging from 1 to 10 resembles the evaluation system that is used in Dutch schools. Third, since the completion frequencies for word stems differ between the English and the Dutch language, it was not possible to translate the English words. In addition, the English word lists contain words with which young children are not familiar. Therefore, the Dutch word lists are composed of other words. These words were chosen by selecting all three letter word stems from a list of word stems with completion frequencies corresponding to those of the English version (Phaf & Wolters, 1991). Then, for each word stem, two words were selected with which the word stem had been completed. The completions with highest and lowest frequencies were not used in order to prevent that correct (or incorrect) responses would depend on completion frequency, rather than on memory performance. Subsequently, we checked in a vocabulary word list for 6 year olds whether the selected words may be considered as part of their passive lexicon (Kohnstamm et al., 1981). This resulted in a list of 60 suitable words that was subdivided into two halves. The two word lists were matched for completion frequency (Phaf & Wolters, 1991), word frequency over a total of one million (Baayen et al., 1995), and word length.

Paired samples $t$-tests revealed that there were no significant differences between the word lists on any of the variables (Completion frequency: $t(29) = -1.21$, $p = .24$; Word frequency: $t(29) = .02$, $p = .98$; Word length: $t(29) = .39$, $p = .69$).
2.4. STATISTICAL ANALYSIS

All analyses were conducted using SPSS (version 16.0) for Windows. For all tests, \( p \) values less than .05 were accepted as significant. To investigate the possible influence of test order and fatigue, the composition of the WCMT word lists and the WMT version, separate univariate analyses (independent \( t \)-tests) were performed for normally distributed variables. Non-parametric statistics (separate Mann-Whitney tests) were used for variables that were not normally distributed. Mann-Whitney tests were also used to test whether the performance of children was comparable to that of healthy adults in case raw adult comparison data were available. If these data were not available, 95% confidence intervals of children’s mean SVT scores were computed to explore the comparability of test scores. Pearson’s \( r \) correlations were used to document associations among SVT test scores and performance on standard neuropsychological tests with normally distributed scores. Where bivariate comparisons involved a non-normally distributed variable, Spearman’s rho rank-ordered correlations were calculated.

To examine possible relationships between the SVTs and cognitive abilities, multiple linear regression analysis was performed on normally distributed variables, whereas logistic regression analysis was performed on non-normally distributed variables. In both approaches, the SVT scores served as the dependent and the cognitive abilities as the independent variables. To reduce the number of variables, test scores on attention and concentration tasks were combined into a summary domain score. The raw scores of the Trail Making Test part A, part B and the total Digit Span of the WISC were transformed into \( z \)-scores, which were subsequently transformed into mean \( z \)-scores. For both regression analyses we used a stepwise procedure, which may capitalize on chance fluctuations in the data. Therefore, also partial correlation was used to explore the relationships between the SVTs and each neuropsychological test, partialling out the effect of the other tests.
3. RESULTS
Of the 48 children that were initially selected, three dropped out. Two did not want to participate, and one did not get permission from his parents. They were replaced by three other randomly selected children.

3.1. CONTROL FOR THE EFFECTS OF TEST ORDER, MATERIAL AND MODALITY
There were no significant differences between the ASTM test scores of children who started or ended the test session with this task ($U = 280.50, p = .88$). Similarly, there were no significant differences between the two groups on the three WCMT measures ($I$ score: $t(46) = -51, p = .61$; $E$ score: $U = 259.00, p = .54$; $R$ score: $t(46) = -.33, p = .74$).

Counterbalancing the two word lists across the Inclusion and Exclusion subtests neither resulted in significant differences on WCMT scores ($I$ score: $t(46) = -1.50, p = .14$; $E$ score: $U = 230.50, p = .23$; $R$ score: $t(46) = -.94, p = .35$).

Using the Bonferroni correction, we found no significant differences between scores on the computerized and orally administered WMT on any of the subtests (WMT IR: $U = 121.00, p = .02$; WMT DR: $U = 195.50, p = .60$; WMT CNS: $U = 143.50, p = .08$; WMT MC: $U = 133.00, p = .04$; WMT PA: $U = 162.00, p = .19$; WMT FR: $t(46) = 1.57, p = .13$).

3.2. CAPABILITY OF CHILDREN TO PASS THE ADULT CRITERIA OF SVTS
Figures 1 to 4 show the scatter plots of the scores on the four SVTs against the children’s age. All children scored above the established cut-offs for adults on the TOMM (Trial 2 and Retention Trial). Similarly, all children passed the WMT using the adult criteria. Five children (10%) failed the ASTM test, all of them being younger than 9 years old. At the WCMT, all children passed according to the cut-off score as defined in this study. However, using the original criterion of this test, 17 of the 48 children (35%) failed.

3.3. COMPARABILITY OF CHILDREN’S AND ADULTS’ SVT PERFORMANCE
In the absence of raw adult comparison data, qualitative analyses showed that children’s scores on Trial 2 and the Retention Trial of the TOMM were practically comparable to those of normal adult controls (Tombaugh, 1997). Children’s WMT scores were tested against available data from 155 normal adults ranging in age from 20 to 80 years ($M = 46.45, SD = 17.78$; Rienstra et al., 2009). For all WMT subtests, we found significant differences between groups, with adults outperforming children (IR: $Z = -5.08, p < .001$; DR: $Z = -2.19, p = .028$; CNS: $Z = -3.78, p < .001$). Likewise, on the
ASTM test, the mean score of children was significantly lower than that of 125 control participants younger than 65 years of age ($M = 32.9, SD = 15.2$; Schmand & Lindeboom, 2004; $Z = -5.49, p < .001$). Finally, a qualitative comparison of the children’s WCMT $R$ scores with those of 69 undergraduates (Hilsabeck et al., 2001) revealed a remarkable difference, in favour of the adults.

FIGURE 1
Distribution Test of Memory Malingering (TOMM) scores by age

[Diagram showing TOMM test scores by age]
FIGURE 2
Distribution Word Memory Test (WMT) subtest scores by age
FIGURE 3
Distribution Amsterdam Short Term Memory (ASTM) test scores by age.
Exploring the value of improved neuropsychological examination

FIGURE 4
Distribution Word Completion Memory Test (WCMT) scores by age

WCMT R score > 9
WCMT R score ≥ 0
Cut-off R score ≤ 0
### 3.4. INFLUENCE OF COGNITIVE IMMATURES ON SVT PERFORMANCE

Correlation coefficients between the effort measures and the neuropsychological tests are presented in Table 1. Since Trial 2 of the TOMM is a constant value, no calculations could be made for this subtest. The Retention Trial of the TOMM was unrelated to any of the cognitive tests. Similarly, for the WMT IR and CNS subtests no significant associations were found. The scores on the DR subtest of the WMT, however, were significantly correlated to all ability tests except Trail Making Test part A. For the ASTM test, significant relationships were found with the letter fluency task, WISC total Digit Span and the 3-MT. The WCMT R score was also positively related to these cognitive tests and, in addition, negatively to Trail Making Test part B.

#### TABLE 1

<table>
<thead>
<tr>
<th></th>
<th>COWAT</th>
<th>TMT part A</th>
<th>TMT part B</th>
<th>WISC Total Digit Span</th>
<th>WISC Vocabulary</th>
<th>3-MT</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOMM Trial 1</td>
<td>.42**</td>
<td>-.27</td>
<td>-.35*</td>
<td>.15</td>
<td>.31</td>
<td>.21</td>
</tr>
<tr>
<td>TOMM Retention a</td>
<td>.26</td>
<td>.10</td>
<td>.09</td>
<td>.02</td>
<td>.00</td>
<td>.00</td>
</tr>
<tr>
<td>WMT IR a</td>
<td>.04</td>
<td>-.05</td>
<td>.00</td>
<td>.03</td>
<td>.16</td>
<td>.17</td>
</tr>
<tr>
<td>WMT DR a</td>
<td>.42**</td>
<td>-.10</td>
<td>-.47**</td>
<td>.41**</td>
<td>.54**</td>
<td>.49**</td>
</tr>
<tr>
<td>ASTM a</td>
<td>.39**</td>
<td>-.20</td>
<td>-.25</td>
<td>.31*</td>
<td>.24</td>
<td>.46**</td>
</tr>
<tr>
<td>WCMT R score b</td>
<td>.53**</td>
<td>-.27</td>
<td>-.30*</td>
<td>.34*</td>
<td>.26</td>
<td>.30*</td>
</tr>
</tbody>
</table>

Note: TOMM Retention = Test of Memory Malingering Retention Trial; WMT IR = Word Memory Test Immediate Recognition; WMT DR = Word Memory Test Delayed Recognition; WMT CNS = Word Memory Test Consistency; ASTM = Amsterdam Short Term Memory test; WCMT R score = Word Completion Memory Test R score; COWAT = Controlled Oral Word Association Test; TMT = Trail Making Test; WISC = Wechsler Intelligence Scale for Children.

- a Spearman rank correlation.
- b Pearson’s r correlation.

* Correlation is significant at the .05 level (two-tailed).
** Correlation is significant at the .01 level (two-tailed).

For the two SVTs on which children failed, regression analyses were conducted with the SVT score as the dependent variable and the letter fluency task, 3-MT, WISC Vocabulary and the composite measure of attention and concentration as the predictors. Due to the uneven distribution of the sample across the two categories on the ASTM
Exploring the value of improved neuropsychological examination test (only 5 participants failed; 10.4%), there is an a priori probability of correct classification of 89.6%. The stepwise logistic regression analysis, however, showed that inclusion of the 3-MT score into the model significantly increased classification accuracy to 95.8% (model Chi-square = 23.82; df = 1; p = .04). Additional inclusion of the other predictors failed to significantly improve the regression model. Consistent with this result, partial correlation showed the strongest association between the ASTM test and the 3-MT, controlling for COWAT and WISC Total Digit Span (r = .61).

To assess possible relationships between the WCMT R score and performance on cognitive tests a multiple linear regression model was fitted. The results indicate that only performance on the letter fluency task made a significant contribution to the prediction of the WCMT R score (F(4; 43) = 4.54, p = .004). The letter fluency test score explained 29% of the WCMT score variance. Correspondingly, the strongest partial correlation (r = .40) was found between the WCMT R score and COWAT, controlling for WISC Total Digit Span, TMT part B and 3-MT.

4. DISCUSSION
This study aimed to provide further evidence for the validity (specificity) of effort tests, by showing that children (whose cognitive skills are not yet fully developed) can pass SVTs.

First, we examined whether young children are able to pass the adult criteria of SVTs. In accordance with our hypotheses, the results indicated that all children passed the established adult cut-off for TOMM Trial 2 and the Retention Trial. This is also consistent with earlier reports that children as young as 5 years of age are able to pass this test (Blaskewitz et al., 2008; Constantinou & McCaffrey, 2003; Donders, 2005; Nagle et al., 2006; MacAllister et al., 2009). In fact, all children in our study achieved the maximum score on Trial 2. As expected, and in line with previous studies (Flaro et al., 2000; Green & Flaro, 2003), all children were also able to pass the WMT effort subtests according to the adult criteria. On the ASTM test, five children failed, all of them being younger than 9 years of age. On the WCMT, all children passed our criterion of an R score equal to or higher than zero. However, using the original cut-off for R scores as defined by Hilsabeck et al. (2001), 17 children failed.

Our second question was whether the SVT performance of children would be comparable to that of healthy adults. The results illustrated that children’s performance on Trial 2 and the Retention Trial of the TOMM was at least as good as that of adult normal controls. Contrary to expectations, the children’s scores on the WMT
subtests were lower than those of healthy adults. This contrasts with previous research that showed that children produced WMT effort scores resembling those of adults (Green & Flaro, 2003). A possible explanation for this discrepancy may be that Green and Flaro (2003) included adult patients with head injuries and a variety of neurological conditions, whereas the present investigation only used healthy normal controls as a comparison group. In fact, the distribution of WMT scores does suggest a considerable influence of age, especially in the youngest children. Green’s automated interpretation program for the WMT (Green et al., 1996) would have given several of the children younger than 9 years old a “caution” based on their performance. This would support earlier findings of Courtney et al. (2003) that performance on the WMT effort measures is significantly related to age.

Furthermore, on the ASTM test, as well as on the WCMT, the mean scores of children turned out to be significantly lower than those of adults. For these two tests, we investigated which cognitive immaturities might be responsible for the reduced level of performance. Regression analyses showed that reading level was the most important predictor of ASTM test performance. Performance on the letter fluency task accounted for a significant proportion of the variance in performance on the WCMT.

Overall, this study adds to the validation of several frequently used SVTs. Children of early school age can easily pass the cut-off scores for sufficient effort of the TOMM and the WMT. Therefore we reasoned that these measures might be relatively insensitive to mild cognitive disorders at an adult age, a conclusion for which there is ample evidence (e.g., Allen & Green, 1999; Iverson et al., 1999; Schagen et al., 1997; Tombaugh, 1997). Furthermore, although the findings in the current study suggest that reading ability significantly contributes to the ASTM test performance, all children older than 9 years passed the test. Thus, only some basic reading skills are required, which adults with schooling higher than grade 4 are expected to possess. For the WCMT, we conclude that mild cognitive impairments should not interfere with obtaining at least an $R$ score equal to zero. Failure according to Hilsabeck’s adult criterion is probably related to immature (in children) or impaired (in adults) verbal fluency.

Because our ‘child-model’ results indicated that there were no significant relations between the SVTs and tests of attention and concentration, mild attention and concentration deficits should not impede successful completion of these SVTs. This is consistent with our hypothesis and it supports earlier findings in adult neurological patients that correlations of SVTs with attention and concentration tasks were low or did not achieve statistical significance (Merten et al., 2007). In addition, the fact that
no effects of test order have been found also corroborates this conclusion. That is, a test session of approximately one hour and a half apparently did not diminish the attentional resources of the children to the extent that it interfered with their performance on SVTs.

In evaluating the present findings, several potential limitations need to be mentioned. First, the number of children who fail the SVTs is small. Therefore, conclusions on the cognitive immaturities that might be responsible for failure on SVTs are based on only a few participants. A second limitation is that a group of children with bona fide neurological impairment, such as children with moderate-severe traumatic brain injury or survivors of cerebellar tumors was not included. Consequently, this study does not add to knowledge base of how valid all these SVTs are in clinical pediatric samples. Finally, an important methodological issue regarding the WCMT needs to be mentioned. As described above, a Dutch translation of the original English WCMT was developed for this study. As yet, there are no data on validity and reliability and neither are there appropriate norms for the Dutch translation. This makes a comparison between the two language versions disputable. Future research should investigate the equivalence of the English and Dutch versions, and validation of the Dutch WCMT is warranted.

In conclusion, our study provides additional support for the validity of several SVTs that are frequently used in clinical practice. The results show that immature cognitive abilities do not interfere with performance on the TOMM and the WMT. The ASTM test requires some basic reading skills. Yet, adults with basic education are expected to possess this level. Additional research on the Dutch WCMT is needed before specific statements can be made. For now, it may be assumed that mild cognitive impairments should not impede remembering at least an equal number of words in each test condition (i.e., to obtain at least an $R$ score equal to zero).

Finally, it needs to be stressed that the above-mentioned conclusions probably do not hold for severe cognitive disorders and clinically obvious symptoms. From previous work, it is known that many patients with severe cognitive impairment fail on SVTs, despite the exertion of optimal effort (Merten et al., 2007). Therefore, more research will be necessary to investigate how performance on SVTs is related to cognitive functioning in order to minimize the chances that patients with genuine cognitive impairments are wrongly designated as suboptimal performers.
Funding

Supported by Psychology Research Institute University of Amsterdam.

Acknowledgement

The authors are grateful to all participating children, to their parents and to the teachers of the Bloemendaalse School Vereniging, Bloemendaal, and the Franciscus Xaverius school, Haarlem, the Netherlands, who supported this study. They also thank dr. P. Green for free-of-charge use of the WMT program and dr. R. Hilsabeck for the permission to translate the English WCMT and for providing the original version of this test.

References


Exploring the value of improved neuropsychological examination.


Differential diagnosis in the memory clinic:


