Validity of an Adapted Scoring Method for a Modified Six Elements Test in Individuals with Brain Injury

Dirk Bertens1,*, Nikita Frankenmolen1, Danielle H.E. Boelen2, Roy P.C. Kessels1,3, Luciano Fasotti1,2,4

1Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen 6500 HE, The Netherlands
2Rehabilitation Medical Centre Groot Klimmendaal, Arnhem 6813 GG, The Netherlands
3Department of Medical Psychology, Radboud University Medical Center, Nijmegen 6525 GA, The Netherlands
4SIZA Support and Rehabilitation, Arnhem 6816 RP, The Netherlands

*Corresponding author at: Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen 6500 HE, The Netherlands.
Tel.: +31-24-3655935; Fax: +31-24-3616066.
E-mail address: d.bertens@donders.ru.nl; dirkbertens.np@gmail.com (D. Bertens)

Accepted 12 December 2014

Abstract

The Modified Six Elements Test (MSET) assesses several executive functions. We examined whether an adapted scoring method is appropriate for discriminating between brain-injured persons with and without executive deficits. A MSET was administered to 70 participants with acquired brain injury in the chronic phase. The group was divided into individuals with and without executive impairments based on several other executive tests. The discriminative value for both the conventional raw score and the adapted scoring method was evaluated using receiver operating characteristic analyses. Both scoring methods discriminated significantly between persons with impaired and unimpaired executive functions (raw score: area under the curve, $AUC = 0.703$, $p = .004$; adapted score: $AUC = 0.780$, $p = .000$). Only the adapted scoring method proved sensitive (81%) and specific (67%) within a clinically useful range. Within this range, an acceptable cut-off score could be determined. Altogether, the proposed MSET scoring index is a potentially clinically useful contribution to the measurement of executive functions.

Keywords: Neuropsychological assessment; Executive functions; Modified Six Elements Test; Acquired brain injury

Introduction

Brain-injured persons with executive deficits often have difficulties in formulating and achieving goals due to deficits in planning and strategy application (Damasio, 1995; Duncan, Emstle, Williams, Johnson, & Freer, 1996; Levine et al., 2000; Lezak, 1982; Stuss & Levine, 2002). As a consequence, unusual and unstructured situations become particularly difficult to handle. However, most neuropsychological tests consist of structured and closed tasks, making reliable assessment of executive functions challenging. A widely used test to assess executive functions is the Behavioural Assessment of the Dysexecutive Syndrome (BADS; Wilson, Alderman, Burgess, Emstle, & Evans, 1996). The aim of this test battery is to assess executive problems in a way that simulates aspects of daily living, by using unstructured and open tasks. The BADS consists of six subtests, measuring different executive processes. One of these subtests is the Modified Six Elements Test (MSET), a measure of multi-tasking ability in which multiple skills such as planning, working memory, prospective memory, rule learning, strategy application, and response monitoring are involved. The MSET has proved to be a good predictor for problems with planning and goal-directed behavior (Alderman, Burgess, Knight, & Henman, 2003; Burgess, Alderman, Evans, Emstle, & Wilson, 1998) and has consistently been found to be one of the most sensitive subtests of the BADS (Bennett, Ong, & Ponsford, 2005; Burgess et al., 1998). Renison, Ponsford, Testa, Richardson, and Brownfield (2012) found that the MSET could reliably predict everyday executive performance in individuals with traumatic brain injury, suggesting that the MSET is an ecologically valid tool. Everyday executive difficulties were measured with the Dysexecutive Questionnaire (Wilson et al., 1996), a 20-item checklist in which cognitive, behavioral, and emotional aspects of executive difficulties were reported by both participants and
D. Bertens et al / Archives of Clinical Neuropsychology 30 (2015); 122–129

In a recent study, Emmanouel, Kessels, Mouza, and Fasotti (2014) showed that the MSET was the most sensitive subtest of the BADS in discriminating between individuals with anterior and posterior lesions. The MSET has been used in various patient groups with different brain disorders in order to evaluate executive functions, for example, persons with MCI and Alzheimer’s dementia (Espinosa et al., 2009), post-concussive symptoms (Chan, Hoosain, Lee, Fan, & Fong, 2003), substance abuse (Fernandez-Serrano, Perez-Garcia, Schmidt Rio-Valle, & Verdejo-Garcia, 2010), schizophrenia (Liu et al., 2011), and traumatic brain injury (Bennett et al., 2005; Chan & Manly, 2002; Manly, Hawkins, Evans, Woldt, & Robertson, 2002). The MSET (Wilson et al., 1996) consists of three tasks (simple arithmetic, written picture naming, and dictation), each of which consists of two parts (subtasks A and B). Participants are instructed that they have 10 min to work on at least a part of each of these six subtasks. However, they are also told that it is not allowed to switch between two parts (subtasks) of the same task. Thus, a participant is not allowed to switch from subtask A of the arithmetic task directly to subtask B of arithmetic, as first one of the other tasks, written picture naming or dictation, has to be dealt with (Burgess et al., 1998). This rule requires participants to engage in task switching (i.e., multitasking) throughout the test. An advantage of the MSET compared with other conventional executive function tests is its highly unstructured character. Therefore, it requires a considerable amount of planning and monitoring behavior. However, the MSET has some limitations as well. Raw scores have a limited range from 0 to 6 and the standardized profile scores, obtained in accordance with the manual of the BADS, range from 0 to 4. In previous studies using the MSET, brain-injured persons obtained mean profile scores of 3 or more (Gouveia, Brucki, Malheiros, & Bueno, 2007; Manly et al., 2002). Individuals with mild executive deficits easily perform at maximum level, indicating that mild executive impairments are not detected by the MSET due to ceiling effects. Furthermore, the dictation task of the MSET requires the use of a tape recorder to record a personal story of the participant. Conventional tape recorders are not only becoming increasingly obsolete, participants have also reported to feel embarrassed by talking about personal topics during this task (Gouveia et al., 2007).

In a recent study (Bertens, Fasotti, Egger, Boelen, & Kessels, 2014), the reliability of an adapted version of the MSET was examined in 60 healthy participants. In the MSET, the dictation subtask was replaced by a sorting task. Moreover, a revised scoring method was proposed, in which the distribution of time spent on the subtasks was taken into account as well. Also, a parallel version of the MSET was added, allowing the measurement of executive functions over the course of time. The test–retest- and parallel form reliability were found to be adequate. However, the validity and diagnostic utility of the adapted scoring method in brain-injured individuals has not been investigated yet.

Here, we examine the diagnostic utility of a recently proposed scoring method (Bertens et al., 2014) of an MSET in persons with brain injuries. Moreover, the adapted scoring method will be compared with the conventional scoring method as proposed in the BADS (Wilson et al., 1996). We also examine to what extent these scoring methods discriminate between individuals with and without executive deficits, based on their performance on other well-established executive tests. Although this comparison does not investigate the test’s ecological validity, that is, its relation with everyday executive problems, our study allows a direct comparison of the sensitivity and specificity of both scoring methods. Furthermore, with the newly proposed scoring method, we aim to reduce ceiling effects, a highly relevant issue in clinical practice. In addition, the convergent validity of both scoring methods will be examined by calculating their correlations with other executive tests.

Methods

Participants

The data of the seventy participants who were included in the current study were collected as part of the recruitment procedure for a larger treatment study, approved by the Medical Review Ethics Committee region Arnhem-Nijmegen (Bertens, Fasotti, Boelen, & Kessels, 2013). All participants had an acquired brain injury of non-progressive nature, such as traumatic brain injury, stroke, or hypoxia. Minimal time since onset of the injury was 3 months. Other eligibility criteria were: age between 18 and 70 years, living independently at home and being fluent in the Dutch language. Participants were recruited from the outpatient department of the Rehabilitation Medical Centre Groot Klimmendaal in Arnhem, the Netherlands and the outpatient rehabilitation clinic for brain-injured individuals and the department of Neurorehabilitation of the Sint Maartenskliniek in Nijmegen, the Netherlands. Exclusion criteria were the presence of severe psychiatric problems; substance abuse (current or in the past); neurodegenerative disorders; severe cognitive comorbidity (e.g., aphasia, dementia). Recruitment was based on assessment by clinicians of the participating centers who excluded all potential participants not meeting the above-mentioned criteria. Executive functions were assessed using six other traditional executive tests.

Table 1 shows the demographic characteristics of the participants. All participants were Caucasian. Level of education was rated using seven categories based on the Dutch educational system, ranging from 1 (less than primary school) to 7 (university degree) (Duits & Kessels, in press). For descriptive purposes, these levels of education were converted to years of education, as is customary in the Anglo-Saxon world (see also Hochstenbach, Mulder, Van Limbeek, Donders, & Schoonderwaldt, 1998).
Neuropsychological Tests

The conventional MSET, a subtest of the executive test battery BADS consists of three open-ended tasks, namely simple arithmetic, dictation, and written picture naming. Each of these tasks consists of two parts (subtask A and B). Participants are instructed to execute at least a part of each of these six subtasks within 10 min. However, it is not allowed to switch between two subtasks of the same type. In the adapted MSET that was used in the present study, the dictation task was replaced by a sorting task in which plastic wall plugs (part A) and small cable ties (part B) had to be sorted by color (three colors) (Bertens et al., 2014). The test was administered according to the instructions of the Dutch version (Krabbendam & Kalff, 1998) of the BADS test manual. The total administration time was ≏ 15 min (5 min for the instructions and 10 min for the execution). For each participant, the number of executed subtasks and the rule breaks were scored, as well as the amount of time spent on each subtask. By subtracting the rule breaks from the number of executed subtask, raw scores were obtained (range 0–6), that were also converted into profile scores (range 0–4) in agreement with the manual of the BADS. In addition, an adapted scoring method was calculated aimed at reducing ceiling effects in mildly impaired individuals (Bertens et al., 2014) using the following formula:

\[
\text{adapted MSET score} = \frac{\text{time longest subtask (s)} - \text{time shortest subtask (s)}}{\text{number of executed subtasks} - \text{rule breaks}}.
\]

If the number of rule breaks is equal to or exceeds the number of executed subtasks, the denominator’s value is set to 1 (i.e., the denominator can never be 0 or a negative number). This scoring index takes into account the distribution of time spent on the six subtasks by subtracting the shortest time spent on one of the six subtasks (i.e., the total time spent on the subtask) from the longest time spent on one of the six subtasks (i.e., the total time spent on the subtask). A more uniform distribution of time across the tasks indicates better multitasking abilities. A lower score indicates better planning performance.

In addition to the MSET, six other neuropsychological tests were administered to assess the main subdomains of executive function (Lezak, Howieson, Bigler, & Tranel, 2012). To assess planning, the Zoo Map test (subtest of BADS) was administered. Response generation was measured with the Category Fluency test and the Letter Fluency test. Response inhibition was assessed with the Go/No-go task from the computerized TAP 2.1 (Zimmermann & Fimm, 2007). The Brixton Spatial Anticipation test (Burgess & Shallice, 1997) was utilized to measure task switching and Letter–Number Sequencing (subtest of WAISIII; Wechsler, 1997) was administered to assess working memory. Scores on the tests were expressed in various standardized scores (i.e., percentile scores, T-scores), in agreement with the respective test manuals. All tests were administered by trained test assistants.

Procedure and Analyses

Based on six executive tests, the participants were divided into an executively impaired group and an executively unimpaired group. The criteria for executive impairments were: a standard score of 1.5 SD below the normative mean on at least two of the executive tests or a standard score between 1 and 1.5 SD below the normative mean on at least four of these tests or a standard score of 1.5 SD below the normative mean on one executive test and a standard score between 1 and 1.5 SD below the normative mean on at least 2 of the remaining executive tests (Bertens et al., 2013).

For statistical analyses, IBM SPSS 20.0 was used. Alpha was set at 0.05 for all analyses and two-tailed tests were used. A multivariate analysis of variance (general linear model) was performed to compare the performance on the MSET between the two groups (executively impaired, executively intact), with both the raw and adapted MSET scores as dependent variables. To compare the demographic characteristics between the groups, t-tests or nonparametric tests for nominal or ordinal variables

### Table 1. Characteristics of the participants with and without executive impairments

<table>
<thead>
<tr>
<th></th>
<th>Executively impaired group</th>
<th>Executively unimpaired group</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>31</td>
<td>39</td>
</tr>
<tr>
<td>Age (mean (SD))</td>
<td>51.7 (± 10.8)</td>
<td>47.5 (± 14.3)</td>
</tr>
<tr>
<td>Sex (N male/female (%))</td>
<td>17/14 (55%/45%)</td>
<td>25/14 (64%/36%)</td>
</tr>
<tr>
<td>Education level (mode (range))**</td>
<td>5 (3–7)</td>
<td>5 (4–7)</td>
</tr>
<tr>
<td>Education (years) (mean (SD))*</td>
<td>11.8 (± 3.1)</td>
<td>14.3 (± 3.4)</td>
</tr>
<tr>
<td>Time past brain injury (months) (mean/median (SD; range))</td>
<td>34.1/15.0 (± 23.1; 4–324)</td>
<td>19.5/11.5 (± 67.0; 3–89)</td>
</tr>
<tr>
<td>Etiology (N (TBI/stroke/other (%))</td>
<td>8/22/1 (26%/71%/3%)</td>
<td>13/24/2 (33%/62%/5%)</td>
</tr>
</tbody>
</table>

Notes: *p < .01; **p < .001.
(sex distribution and education level) were conducted. To control for possible effects of demographic differences, a multivariate analysis of covariance was performed when applicable.

The discriminative value of the MSET for executive deficits was evaluated by calculating receiver operating characteristic (ROC) curves for both the raw score and the adapted score separately. The area under the curve (AUC) indicates the discriminative power of the test, which varies between 0.5 (no discriminative power) and 1.0 (maximum discriminative power). Within this analysis, an optimal cut-off point was determined, fulfilling the criteria of a good sensitivity (>80%) and an acceptable specificity (>60%) rate (see Blake, McKinney, Treece, Lee, & Lincoln, 2002; Kessels, Mimpen, Melis, & Olde Rikkert, 2009; Oosterman, Molenveld, Olde Rikkert, & Kessels, 2010).

To examine the convergent validity of the MSET and the other executive tests, Pearson correlation coefficients (r) were calculated between both the raw and the adapted scores of the MSET and the other executive tests. For this analysis, the scores of the other executive tests were expressed in raw scores, with higher scores indicating better test performance. For this reason, the scores on the Brixton were expressed in number of correct items instead of errors, whereas the scores on the Go/No-go task were multiplied by −1.

Results

Based on the six executive tests, the participants were divided into an executively impaired group (N = 31) and an executively unimpaired group (N = 39). Intergroup comparisons showed no significant difference for age [t(68) = −1.36; p = .18], but a significant difference for education level [U = 336.00; p < .001] and years of education [t(68) = 3.26, p = .002]. No differences between the groups with respect to post onset time [t(42) = −0.97, p = .34] and sex distribution were present [χ²(1) = 0.62; p = .43].

Table 2 shows the MSET scores for both the executively impaired and unimpaired groups. Significant differences on the MSET scores between the executively impaired group and the executively unimpaired group were found [F(2, 67) = 6.30, p = .003], with the executively impaired group performing worse than the executively unimpaired group (i.e., as reflected by a lower raw score and a higher adapted score). These group differences were found for both the raw scores [F(1, 68) = 9.72, p = .003] with a moderate-to-large effect size (Cohen’s d = 0.71) and the adapted MSET scoring method [F(1, 68) = 11.85, p = .001] with a moderate-to-large effect size (d = 0.79). As educational levels differed significantly between the groups, a multivariate analysis of covariance was conducted with education level as covariate. This analysis showed that education level [F(2, 66) = 1.11, p = .34] did not contribute significantly, and the overall group differences remained statistically significant [F(2, 66) = 7.20, p = .001].

The discriminative value of both scoring methods was evaluated using ROC analyses. The ROC graph of the raw MSET score as predictor of executive function deficits showed a significant AUC (AUC = 0.703, p = .004; Fig. 1A). However, no optimal cut-off score with a sensitivity of >80% and a specificity of >60% could be determined (Table 3). Moreover, an ROC curve with an AUC ≤ 0.75 is generally interpreted as clinically not useful (Fan, Upadhye, & Worster, 2006). The ROC analysis of the adapted MSET score, as a predictor for executive deficits, also showed a significant area under the curve (AUC = 0.780, p = .000; Fig. 1B). Moreover, for this scoring index, it was possible to determine an optimal cut-off score (18.63) with a sensitivity of 81% and a specificity of 67%, indicating that a score of 18.63 or more is an indication for executive deficits (Table 4). Statistical comparison between the areas under the ROC curves of both scoring indices (Hanley & McNeil, 1982) did not show a significant difference between the AUCs (p = .35).

With respect to the convergent validity of both scoring methods, an expected high negative correlation was found between the raw and the adapted MSET scores, r = −.745, p < .01, in which higher raw scores and lower adapted scores represent better test performance. Correlations between both scoring methods of the MSET and the other executive tests are reported in Table 5. Both scoring indices showed a similar pattern. That is, for the raw score significant, but moderate (cf. Cohen, 1992) correlations were found with three of six executive function tests. Four of the six tests correlated significantly with the adapted score, also in the moderate range.

### Table 2. Group differences for the raw and adapted scores on the Modified Six Elements Test (MSET)

<table>
<thead>
<tr>
<th></th>
<th>Raw MSET score</th>
<th></th>
<th>Adapted MSET score</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>(SD)</td>
<td>Range</td>
<td>M</td>
</tr>
<tr>
<td>Executively impaired</td>
<td>4.4</td>
<td>(± 1.6)</td>
<td>0–6</td>
<td>52.4</td>
</tr>
<tr>
<td>Executively intact</td>
<td>5.4</td>
<td>(± 1.2)</td>
<td>2–6</td>
<td>21.6</td>
</tr>
</tbody>
</table>

Notes: The raw score (in accordance to the BADS manual) consists of a scale with a range from 0 (lowest) to 6 (highest). The adapted MSET score is a continuous variable in which a lower score indicates better performance.
Finally, we examined the occurrence of ceiling effects by examining the frequency of participants who obtained the maximum score. Table 6 shows the distribution of the raw MSET scores, followed by the statistics of the corresponding adapted MSET scores. Forty-one of seventy participants had a maximum raw score of 6 (equivalent to a maximum profile score of 4), 11 of whom were in the “impaired” group ($n = 31$) and 30 in the “unimpaired” group ($n = 39$). The most frequently found score in both groups was the maximum score. The maximum adapted MSET score (reflecting an equal distribution of time spent on all 6 tasks without any rule

![Fig. 1. Receiver operating characteristic curves for the conventional MSET raw scoring method (A) and for the adapted MSET scoring method (B).](image)

<table>
<thead>
<tr>
<th>Cut-off point</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>13</td>
<td>95</td>
</tr>
<tr>
<td>3</td>
<td>29</td>
<td>85</td>
</tr>
<tr>
<td>4</td>
<td>52</td>
<td>85</td>
</tr>
<tr>
<td>5</td>
<td>64</td>
<td>77</td>
</tr>
</tbody>
</table>

Table 3. Possible cut-off scores for the conventional raw MSET scoring method with the corresponding sensitivity and specificity

<table>
<thead>
<tr>
<th>Cut-off point</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.59</td>
<td>81</td>
<td>61</td>
</tr>
<tr>
<td>18.09</td>
<td>81</td>
<td>64</td>
</tr>
<tr>
<td>18.63*</td>
<td>81</td>
<td>67</td>
</tr>
<tr>
<td>20.04</td>
<td>77</td>
<td>67</td>
</tr>
<tr>
<td>21.67</td>
<td>77</td>
<td>69</td>
</tr>
<tr>
<td>22.50</td>
<td>74</td>
<td>69</td>
</tr>
</tbody>
</table>

Table 4. Possible cut-off scores for the MSET adapted scoring method with the corresponding sensitivity and specificity

<table>
<thead>
<tr>
<th>Cut-off point</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw MSET score</td>
<td>0.246</td>
<td>0.410**</td>
</tr>
<tr>
<td>Adapted MSET score</td>
<td>-0.243</td>
<td>-0.395**</td>
</tr>
</tbody>
</table>

Table 5. Pearson correlations between the MSET raw and profile score and adapted MSET score and other executive function tests

<table>
<thead>
<tr>
<th></th>
<th>Zoo Map</th>
<th>LFT</th>
<th>CFT</th>
<th>LNS</th>
<th>Brixton</th>
<th>Go/No-go</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw MSET score</td>
<td>0.246</td>
<td>0.410**</td>
<td>0.325**</td>
<td>0.463**</td>
<td>0.151</td>
<td>-0.225</td>
</tr>
<tr>
<td>Adapted MSET score</td>
<td>-0.243</td>
<td>-0.395**</td>
<td>-0.262*</td>
<td>-0.504***</td>
<td>-0.146</td>
<td>0.238*</td>
</tr>
</tbody>
</table>

Notes: LFT = Letter Fluency Test; CFT = Category Fluency Test; LNS = Letter–Number Sequencing; Brixton = Brixton Spatial Anticipation Test. *p < .05, **p < .01, ***p < .001 (two tailed).

Finally, we examined the occurrence of ceiling effects by examining the frequency of participants who obtained the maximum score. Table 6 shows the distribution of the raw MSET scores, followed by the statistics of the corresponding adapted MSET scores. Forty-one of seventy participants had a maximum raw score of 6 (equivalent to a maximum profile score of 4), 11 of whom were in the “impaired” group ($n = 31$) and 30 in the “unimpaired” group ($n = 39$). The most frequently found score in both groups was the maximum score. The maximum adapted MSET score (reflecting an equal distribution of time spent on all 6 tasks without any rule
breaks) is 0, a score which was not achieved by any of the participants. For the 41 participants who obtained the maximum raw score, the adapted MSET scores ranged from 2 to 51 (mean score 15.8; SD 10.2) indicating an improved variability and no ceiling effect.

**Discussion**

The aim of this study was to examine the diagnostic utility of an updated version of the MSET (Bertens et al., 2014) including an adapted scoring method. The results show that persons with executive function deficits performed significantly worse on the MSET than persons without executive deficits, regardless of scoring method. With respect to the discriminative value, the AUCs for the raw score and the adapted score were both statistically significant. Although the AUCs did not significantly differ, only the AUC for the adapted score was within a clinically useful range (Fan et al., 2006). An acceptable cut-off score, fulfilling the criteria of having a good sensitivity and an adequate specificity for discriminating between impaired and normal executive function, could only be determined for the adapted score.

An altered version of the MSET was used in which the dictation subtask was replaced by a sorting task, making the test more suitable for individuals with language impairments. Shallice and Burgess (1991) described that the aim of the original SET was to evaluate if a participant could devise a simple plan, scheduling the subtests (consisting of “fairly simple” activities) efficiently and keeping a check on time. The main objective was to evaluate the application of an efficient strategy to alternate between the tasks (switching without rule breaks). Results within subtasks are not taken into account. Therefore, changing a subtask for another simple task should have marginal effect on the construct of the test. Moreover, the dictation task is also replaced by a sorting task in the BADS-C, the children’s version of the BADS (Emslie et al., 2003).

The compatibility between both MSET scoring indices and other executive tests is acceptable. For the adapted scoring method, significant correlations were found with four of six executive tests; however, these correlations were weak to moderate. No correlations were found between the MSET and the Zoo Map test, measuring planning (see also Oosterman, Wijers, & Kessels, 2013) and the Brixton Spatial Anticipation Test, measuring shifting (Van den Berg et al., 2009). Although planning and shifting may seem important aspects, the aim of the MSET was to measure executive functions in an ecologically valid way using unstructured and open-ended tasks which better mimics everyday demands. This may explain the low-to-moderate and even absent correlations between the MSET and traditional more structured executive tests.

A limitation of the current study is that the executive tests used to divide the participants in an executively impaired and executively intact group were traditional executive tests. As a result, we cannot conclude that our revised scoring method also optimizes the test’s ecological validity. One could argue that a comparison with other ecologically valid (e.g., open-ended tasks which resemble real life duties) tasks is required to determine the convergent validity of the MSET. Examples of such tasks are the Executive Secretarial Task (Lamberts, Evans, & Spikman, 2010), the Multiple Errands Test (Alderman et al., 2003), and the Hotel Task (Manly et al., 2002). However, although these tasks are standardized assessment procedures, their psychometric properties (e.g., sensitivity and specificity) have not been examined thoroughly and normative data are not available. More importantly, administration of these tasks is very time-consuming and may last several hours. Therefore, these paradigms are less suitable for implementation in clinical practice (see also Lamberts et al., 2010).

Future studies should further examine which executive aspects the MSET is tapping. Moreover, the ecological validity of the proposed scoring method should be further investigated by including questionnaires such as the Dysexecutive Questionnaire (Wilson et al., 1996) or the Executive Function Index (Spinella, 2005) that assess everyday executive difficulties or by using established everyday executive function tests. The validity of the adapted scoring method should be investigated in other clinical populations as well. Another limitation of our study is that other cognitive domains were not assessed. Therefore, we cannot rule out that

<table>
<thead>
<tr>
<th>Raw MSET score</th>
<th>Score</th>
<th>N</th>
<th>Mean</th>
<th>(SD)</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>139.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>110.1</td>
<td>84.7</td>
<td>37–211</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>64.2</td>
<td>32.7</td>
<td>30–110</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>51.3</td>
<td>27.8</td>
<td>19–93</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>27.5</td>
<td>7.1</td>
<td>21–39</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>41</td>
<td>15.8</td>
<td>10.2</td>
<td>2–51</td>
<td></td>
</tr>
</tbody>
</table>

Table 6. Distribution of the adapted MSET scores compared with the raw MSET scores
the two groups also differed on memory or attention performance. Eventual differences could have influenced the results on the executive function tasks, including the MSET.

To our knowledge, this is the first study investigating the MSET taking the aspect of time distribution into account. The conventional scoring method as described in the manual of the BADS only includes a penalty profile point when individuals spend >271 s on one of the MSET subtasks (Wilson et al., 1996). It lacks a scoring system for unequal time distribution throughout the subtests. In order to equally distribute the allowed time over the six subtests, planning behavior and time monitoring are required (Mantyla, Carelli, & Forman, 2007). In the current study, time distribution was assessed by subtracting the shortest time spent on a subtask from the longest time spent on a subtask. Unfortunately, this method does not take into account the overall variability of time spent on the subtasks. However, it is a relatively simple method to estimate time distribution and feasible in clinical practice. Adding the aspect of time distribution to the scoring index prevents the occurrence of ceiling effects in persons with mild executive deficits. This improvement is indicated by the better sensitivity of the adapted MSET score (resulting in a clinically useful cut-off point) compared with the poor sensitivity of the traditional score. An unintended outcome of this study was that the two investigated groups differed significantly on some demographic characteristics. The impaired group had a lower education level, which may affect performance on the executive tests. However, adjusting for this potentially confounding factor did not alter the results.

In conclusion, the present study shows that the adapted scoring method of the MSET can be clinically useful in measuring executive deficits in individuals with brain injuries, and more sensitive and specific than the conventional score. The adapted scoring method of the MSET is able to discriminate between persons with and without executive function deficits. The next step is to collect normative data in a sample of healthy participants from various age groups and the full spectrum of educational levels, for use in clinical assessment. Although the MSET, including the adapted scoring index, may be clinical useful as a first executive screening or as a treatment outcome measure, it is always advised to use multiple tests for the assessment of executive functions, since executive function is a multifarious concept.

Acknowledgements

The authors acknowledge the participation of the study participants and thank Rehabilitation Medical Centre Groot Klimmendaal and the Sint Maartenskliniek for their contribution to the data collection. This work was supported by the National Initiative Brain and Cognition (NIBC). This quick-result project is embedded in the pillar “The Healthy Brain, Program Cognitive Rehabilitation” (grant number 056-11-011).

Conflicts of Interest

None declared.

References


