Utility of TOMM Trial 1 as an Indicator of Effort in Children and Adolescents†

Brian L. Brooks1,*, Elisabeth M.S. Sherman1, Andrea L. Krol2

1Alberta Children’s Hospital and University of Calgary, Calgary, Alberta, Canada
2University of Alberta, Edmonton, Alberta, Canada

*Corresponding author at: Neurosciences Program, Alberta Children’s Hospital, 2888 Shaganappi Trail NW, Calgary, Alberta, Canada T3B 6A8.
Tel.: 1-403-955-2597; fax: 1-403-955-7501.
E-mail address: brian.brooks@albertahealthservices.ca (B.L. Brooks).
Accepted 30 September 2011

Abstract

Although measuring test compliance in a pediatric neuropsychological evaluation is important, increasing demands on clinicians’ time and the need for efficiency during assessments may make it difficult to routinely include effort testing. This study investigated whether performance on Trial 1 of the Test of Memory Malingering (TOMM) is predictive of overall performance in children and adolescents with neurological disorders. Participants included 53 children and adolescents between six and 19 years (mean age = 12.4, SD = 4.1) who were followed through a neurology clinic at a tertiary care hospital. Several cutoff scores were examined, with the goal of maximizing positive predictive (accurate detection of failure on the TOMM) and negative predictive (accurate detection of passing the TOMM) values. Every participant who scored ≥36 on Trial 1 (n = 50) went on to pass the TOMM. This study is the first step in providing evidence that performance on Trial 1 might be used as a quick screening measure of overall performance on the TOMM in children and adolescents. Further research on this topic is warranted.

Keywords: Effort; Malingering; Symptom validity; Malingered neurocognitive disorder; Pediatrics; Children

Introduction

The evaluation of effort, or test-taking compliance, is an important component of neuropsychological assessments with children and adolescents. There is increasing evidence that healthy children and adolescents can readily pass effort tests (Constantinou & McCaffrey, 2003; Rienstra, Spaan, & Schmand, 2010), that these measures can be used with individuals with neurological disorders (Carone, 2008; Donders, 2005; Donders & Boonstra, 2007; Green & Flaro, 2003; Green, Flaro, & Courtney, 2009; Kirk et al., 2011; Kirkwood & Kirk, 2010; MacAllister, Nakhtina, Bender, Karantzoulis, & Carlson, 2009), and that they can identify children and adolescents who are asked to simulate cognitive impairments (Blaskewitz, Merten, & Kathmann, 2008; Frazier, Frazier, Busch, Kerwood, & Demaree, 2008; Gunn, Batchelor, & Jones, 2010; Nagle, Everhart, Durham, McCammon, & Walker, 2006). There is, of course, some caution needed when interpreting effort test performance in those who are very young or have lower reading and/or cognitive skills (e.g., Courtney, Dinkins, Allen, & Kuroski, 2003), but the literature supports the utility of effort tests for children as young as 5 years of age (i.e., the reader is referred to Kirkwood [in press] for an eloquent summary of the current research on effort testing in children and adolescents).

One effort test, in particular, the Test of Memory Malingering (TOMM; Tombaugh, 1996), has received particular attention regarding its use as an effort test in children and adolescents. Healthy children and adolescents are readily able to pass the TOMM (i.e., failure rates range from 0% [Blaskewitz et al., 2008; Constantinou & McCaffrey, 2003; Rienstra et al., 2010]

† These data were presented at the 31st annual meeting of the National Academy of Neuropsychology, Marco Island, FL.
to 2% \citep{Gunn2010}. The vast majority of children and adolescents with medical or neurological diagnoses are also able to pass the TOMM (i.e., failure rates from the literature have included 3% \citep{Donders2005} and 4% \citep{Kirk2011} in mixed clinical samples and 10% in a pediatric epilepsy sample \citep{MacAllister2009}). The TOMM has also shown promise for detecting intentional suboptimal effort in children and adolescents. For example, the correct identification of simulated malingerers ranges from 68\% \citep{Blaskewitz2008} to 95\% \citep{Gunn2010} accuracy.

Despite increasing support of the TOMM for detecting poor effort in children and adolescents, some clinicians might not routinely assess for effort because of limited time and resources. One potential solution in non-forensic settings is to administer an abbreviated version of the TOMM and use this information as an indicator of overall performance (i.e., the TOMM is comprised of three trials, but an abbreviated version could include only the first trial). In adults, several studies have indicated that examinees who pass Trial 1 of the TOMM also pass the entire TOMM \citep{Bauer2007, Gavett2005, Horner2006, O'Bryant2007, O'Bryant2008}. However, there has yet to be a study examining whether performance on the first trial is indicative of overall TOMM performance in children and adolescents.

As clinicians search for more time-efficient assessment methods, there is increasing need to examine the clinical utility of shorter versions of existing measures. The purpose of this study was to examine whether performance on Trial 1 is indicative of overall performance on the TOMM in children and adolescents. Because children’s performance on the full version of the TOMM is similar to that of adults’ performance on the TOMM \citep{Kirkwood2011}, it is hypothesized that Trial 1 of the TOMM will be predictive of overall TOMM performance in children, consistent with the adult literature.

**Methods**

**Participants**

Participants included 53 children and adolescents between the ages of 6 and 19 (mean = 12.4, SD = 4.1) who were referred for a neuropsychological assessment through a neurology clinic at a tertiary care hospital. Inclusion criteria for this study were a formal neurological diagnosis provided by the neurologist and the ability to complete testing.

**Measures**

The TOMM \citep{Tombaugh1996} is a forced-choice, visual recognition test. The target drawings are presented one at a time, followed by forced-choice recognition (Trial 1). This process is repeated for a second learning trial of the same information (Trial 2). After a 15-min delay, if performance on Trial 2 was below the established cutoff score, then the examinee is again asked to pick out the learned information through forced-choice recognition (Retention Trial).

The children and adolescents in this study were administered the TOMM according to the instructions set forth in the manual and the cutoff established in previous adult literature. Based on the TOMM manual, the retention trial is optional when a participant’s scores above the established cutoff on Trial 2. The TOMM was included as part of a larger neuropsychological assessment. Other measures also used in these assessments included the Full-Scale IQ (FSIQ) from the Wechsler Intelligence Scale for Children-Fourth Edition \citep{WISCIV, Wechsler2003} or the Wechsler Adult Intelligence Scale-Third/Fourth Edition \citep{WAISIII/IV, Wechsler1997a, Wechsler2008} for overall intelligence, the Continuous Attention Test \citep{Seidel1991} for sustained attention, the Long-Delay Free Recall score from the California Verbal Learning Test (CVLT)-Children’s Version \citep{Delis1994} or CVLT-Second Edition \citep{Delis2000} for verbal memory, the Faces Delayed subtest from the Children’s Memory Scale \citep{Cohen1997b} or Wechsler Memory Scale, Third Edition \citep{Wechsler1997b} for visual memory, the Behavior Rating Inventory of Executive Function \citep{Gioia2000} as a measure of executive functioning, and the Scales of Independent Behavior-Revised \citep{Bruininks1996} as a measure of adaptive functioning. Information on the other neuropsychological measures was included in order to provide a better description of the cognitive abilities of this sample.

**Analyses**

The accuracy level for several cutoff scores on Trial 1 was examined. However, it is important to note that for the purposes of this study, an overall “pass” score was based on adequate performance on Trial 2 or Retention, consistent with the established cutoff scores set forth in the TOMM administration manual \citep{Tombaugh1996} and prior pediatric literature \citep{Donders2005, Kirk2011, MacAllister2009}. 


Several indicators of classification accuracy were explored in this study, including sensitivity (SN), specificity (SP), positive predictive value (PPV), and negative predictive value (NPV). SN is the proportion of persons who (a) fail Trial 1 and fail overall on the TOMM (true positive, TP) to (b) those who pass Trial 1 but fail overall on the TOMM (false negative, FN). SP is the proportion of persons who (a) pass Trial 1 and pass overall on the TOMM (true negative, TN) to (b) those who fail Trial 1 but pass overall on the TOMM (false positive, FP). Related to SN and SP are PPV (e.g., the proportion of persons who are identified as failing the TOMM based on Trial 1, in relation to all persons who do poorly on Trial 1) and NPV (e.g., the proportion of persons who are identified as passing the TOMM based on Trial 1, in relation to all persons who pass Trial 1). PPV and NPV also take into consideration the base rates (BRs) of poor performance on the TOMM, with the following formula (see below) and the BRs of 3% (Donders, 2005), 4% (Kirk et al., 2011), and 10% (MacAllister et al., 2009). For this study, maximizing PPV and NPV was desired.

\[
\begin{align*}
SN &= \frac{TP}{TP + FN}, \\
SP &= \frac{TN}{TN + FP}, \\
PPV &= \frac{SN \times BR}{(SN \times BR) + ((1 - SP) \times (1 - BR))}, \\
NPV &= \frac{SP \times (1 - BR)}{(SP \times RC) + ((1 - SN) \times BR)}
\end{align*}
\]

**Results**

Table 1 presents the demographic information for this sample. The mean age of this sample was 12.4 (SD = 4.1) and all participants were between 6 and 19 years of age. This was a mostly Caucasian (71.7%) but diagnostically heterogeneous group (epilepsy \(n = 24; 45.2\%\), traumatic brain injury \(n = 9; 16.9\%;\) with one patient sustaining numerous concussions in hockey and eight patients sustaining moderate to severe brain injuries], stroke \(n = 7; 13.2\%\), hydrocephalus \(n = 9; 16.9\%), and other \(n = 4; 7.5\%;\) including neurofibromatosis-1, Behcet’s disease, autoimmune encephalitis, and craniopharyngioma).

Table 2 presents the cognitive abilities and overall TOMM scores for this pediatric neurology sample. Mean intellectual abilities on the WISC-IV FSIQ were >1 SD below healthy children (mean = 83.6, SD = 18.8), with over one-quarter falling within the “impaired” range. When considering those children who were at least 2 SD below the mean, and thus considered to have impaired abilities, the percentages ranged from 13.6% with visual memory to 50.0% with sustained attention. Mean scores for TOMM trials were 44.0 (SD = 5.6) for Trial 1, 48.4 (SD = 5.0) for Trial 2, and 47.8 (SD = 6.4) for Retention. On Trial 1, 41.5% of the sample scored below the established cutoff set forth in the manual. The overall failure rate for the TOMM, based on performance on Trial 2, was 5.7%. For the retention trial, the failure rate went up slightly because of one child who passed Trial 2 but scored 1 point below the cutoff on the retention trial.

The accuracy of Trial 1 compared with overall performance on the TOMM is presented in Table 3. This information is based on different cutoff scores on Trial 1 (ranging from 30 to 45). Trial 1 scores of ≥36 resulted in 100% NPV. In other words, if a

<p>| Table 1. Demographics of pediatric neurology patients |
|-----------------|-----------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>Demographics</th>
<th>(N)</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>53</td>
<td>12.4</td>
<td>4.1</td>
<td>6–19</td>
</tr>
<tr>
<td>Parent Education (years)</td>
<td>52</td>
<td>13.3</td>
<td>2.3</td>
<td>8.5–20</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Percentage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>38</td>
<td>71.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>0</td>
<td>0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>2</td>
<td>3.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>8</td>
<td>15.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>5</td>
<td>9.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnosis</td>
<td>Percentage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epilepsy</td>
<td>24</td>
<td>45.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traumatic Brain Injury</td>
<td>9</td>
<td>16.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroke</td>
<td>7</td>
<td>13.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hydrocephalus</td>
<td>9</td>
<td>16.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>7.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: SD = standard deviation.
child with a neurological condition score of ≥36 on Trial 1, then there is a 100% chance that the child will pass the TOMM. Trial 1 scores of ≤34 resulted in an SP of 0.98 but low PPVs (0.79 with 10% BR of failure, based on MacAllister et al. [2009], 0.58 with a 4% BR, as derived from Kirk et al. [2011], and 0.51 with a 3% BR, as derived from Donders [2005]). When
examining scores of \( \leq 33 \), SP was 1.0 and PPV was 1.0 (i.e., all three children who scored \( \leq 33 \) on Trial 1 did not pass the TOMM). The SN, SP, PP, and NPV for various cutoff scores and different BRs (derived from the current studies of the TOMM in pediatric samples) are presented in Table 3 for clinicians’ use.

Discussion

The purpose of this study was to determine whether performance on the first trial of the TOMM is indicative of overall performance on the TOMM in children and adolescents. It should be noted that this study was not designed to examine the validity of Trial 1 from the TOMM as an effort test in pediatrics, to determine whether the TOMM is the most appropriate measure of effort for pediatrics, to indicate that more extensive effort testing is not warranted in pediatric neuropsychological assessments, or that the full version of the TOMM should not be administered. This study does, however, provide initial support for the use of the first trial of the TOMM in those clinical (and research) situations where effort testing might not normally occur due to time demands or other factors that limit the ability to administer the full TOMM.

When considering Trial 1 as an indicator of overall TOMM performance, this study provides preliminary evidence that performance on Trial 1 provides a good estimate of overall TOMM performance. In those children and adolescents who passed Trial 1 (based on the cutoff established in the test manual), 100% went on to pass the full TOMM. This level of predictive accuracy is consistent with the adult studies that have examined Trial 1 performance in relation to overall TOMM performance (Bauer et al., 2007; Gavett et al., 2005; Horner et al., 2006; O’Bryant et al., 2007, 2008). Perhaps more importantly, 100% of children and adolescents who obtain a score of \( \geq 36 \) on Trial 1 also pass the TOMM. The broad range of Trial 1 scores obtained in patients who still passed the TOMM might be related to the characteristics of this clinical sample. In an adult study of Trial 1 utility, for example, a 100% level of accuracy for detecting failure on the TOMM was only reported for cutoff scores of \( \geq 44 \) (Bauer et al., 2007). In that same study, Trial 1 scores of \( \geq 36 \) had NPVs ranging from 87.8% to 97.7% (depending on the BR). Of course, older age (mean age = 40.5 years, \( SD = 14.3 \)), higher IQ (mean WAIS FSIQ = 96.4, \( SD = 13.7 \)), larger sample size, and different clinical context (e.g., litigants with mild traumatic brain injuries) in the Bauer and colleagues (2007) study may account for some of the differences compared with the present sample. It would be important for future research to examine Trial 1 accuracy in pediatric patients with mild traumatic brain injuries who are involved in litigation, as the results from Bauer and colleagues (2007) might suggest that there could be different results for these patient characteristics.

The overall failure rate on the TOMM in this mixed pediatric neuropsychology sample was 5.7%. Although the purpose of this study was not to explore the validity of the TOMM for detecting poor effort in children and adolescents, it is noteworthy that the failure rate is consistent with the levels reported in three other pediatric samples (i.e., 3%–10%) where the validity and utility of the TOMM as a measure of effort was the central research question (Donders, 2005; Kirk et al., 2011; MacAllister et al., 2009). Of the three patients who failed the TOMM, it is believed that one of them was an FP due to extremely low cognitive abilities (e.g., 7-year-old with refractory epilepsy, WISC-IV FSIQ = 44, verbal and visual memory < 1st percentile, and parent- and teacher-rated adaptive functioning < 1st percentile). On the other hand, it is possible that the other two patients who failed the TOMM were providing insufficient effort on this measure. The other patients, a 15-year-old boy with neurofibromatosis-1 and a 14-year-old girl with Behcet’s disease and juvenile arthritis, were suspected of being TP cases based on their chance-level performance across the TOMM despite broadly normal cognitive abilities (including memory testing) and ability to pass another measure of effort (i.e., Victoria Symptom Validity Test; Slick, Hopp, & Strauss, 1997). Interestingly, when the male patient was re-tested 1 year later, his performance on the TOMM was above the established cutoff scores. Based on the closer inspection of these three cases who performed below the cutoff, perhaps it is more accurate to say that the failure rate on the TOMM was 3.8%.

There are some limitations to consider for this study. First, the sample size was small and contained a range of ages, diagnoses, and levels of functioning. It is possible that a larger sample size, or patient groups with a more homogeneous presentation, could perform differently on these analyses of the accuracy of Trial 1. Second, the failure rate on the TOMM was low, which resulted in some of the analyses with only a few participants who scored below the cutoffs. Related to the first limitation, it is possible that a larger sample size could result in a different failure rate (although it is relatively consistent with other pediatric studies). The failure rate in this study does, of course, speak to the ability of low functioning children to pass the TOMM. Third, this was a sample of consecutively referred patients evaluated within a tertiary care hospital. As such, the cognitive abilities of this sample might not reflect patients assessed in other neuropsychological settings, where performance on the initial trial might be different. Fourth, this sample was not a forensic sample, and the focus of the neuropsychological assessments was clinical care. It is possible that failure rates on the TOMM and the accuracy of Trial 1 for determining overall performance could differ in a pediatric forensic sample.

This is the first study to suggest that performance on Trial 1 can be indicative of overall TOMM performance in children and adolescents. In this sample of pediatric neurology patients, every patient who scored \( \geq 36 \) on Trial 1 also passed the TOMM. At
this time, however, the results are likely preliminary based on the relatively smaller sample size. Further research is warranted to examine the accuracy of Trial 1 for detecting poor effort in larger neurology samples, as well as in other clinical samples with higher BRs of poor effort.

Conflict of Interest

Drs Brooks and Sherman receive funding from Psychological Assessment Resources, Inc. and book royalties from Oxford University Press.

Acknowledgements

The authors would like to thank Helen Carlson, PhD, Lonna Ivey, Emily Tam, and Julie Wershler for their assistance with data entry and manuscript preparation.

References


