Performance of forensic and non-forensic adult psychiatric inpatients on the Test of Memory Malingering

Susan D. Gierok a,*, Andrew L. Dickson b, Jack A. Cole c

a Taylor Hardin Secure Medical Facility, 1301 Jack Warner Parkway Northeast, Tuscaloosa, AL 35404, USA
b Department of Neurology, Hattiesburg Clinic, 415 South 28th Avenue, Hattiesburg, MS 39401, USA
c Samaritan Center, 515 Bayou Street, Vincennes, IN 47591, USA

Accepted 9 April 2005

Abstract

This study compared performance on the Test of Memory Malingering (TOMM [Tombaugh, T. N. (1996). Test of Memory Malingering (TOMM). New York: Multi Health Systems]) between a Forensic Psychiatric group and a Non-forensic Psychiatric group of 20 men each. It was hypothesized that the Forensic group would perform less well on the TOMM due to greater secondary gain for that population. The Forensic group (age, M = 32.65 years; 16/20 were minorities) was composed of inpatients from a forensic psychiatric facility who had been referred for pre-trial evaluations. The Psychiatric group (age, M = 41.00 years; 15/20 were Caucasian) were chosen from an inpatient psychiatric facility and had no pending legal involvement. As hypothesized, the Psychiatric group performed significantly better than the Forensic group on all TOMM trials. A TOMM score of below 45 on Trial 2 or the Retention Trial is consistent with probable response bias. Only one member of the Psychiatric group (the same individual) met this criterion, whereas seven members of the Forensic group met this criterion. The TOMM identified patients with pending legal charges as more likely to exert less effort than those with no obvious secondary gain.

Keywords: Malingering, TOMM, Forensic psychiatric; Tests of effort

Several researchers have noted the importance of assessing malingering in the medicolegal setting because of potential secondary gain (e.g., Iverson & Binder, 2000). The Test of Memory Malingering (TOMM; Tombaugh, 1996), a 50-item, forced-choice, visual recognition test, was...
used to compare performance between a forensic inpatient psychiatric group (Forensic) and an inpatient psychiatric group (Psychiatric). Two learning trials and a retention trial comprise the TOMM, with feedback provided for each trial (Tombaugh, 1996). A score below 45 on Trial 2 or the Retention Trial is consistent with probable response bias and would contribute to the impression of probable malingering of neurocognitive dysfunction (Slick, Sherman, & Iverson, 1999).

The detection of malingering is a difficult one. Yet, forensic examiners are faced with this task routinely. This study was conducted to examine the usefulness of the TOMM as an instrument to aid in the detection of malingering in forensic psychiatric evaluations. It was hypothesized that significantly lower scores would be demonstrated by the Forensic group because of greater secondary gain (e.g., avoidance of criminal proceedings). The TOMM has demonstrated high levels of both sensitivity and specificity (e.g., Rees, Tombaugh, Gansler, & Moczynski, 1998) and is largely insensitive to depression (Rees, Tombaugh, & Boulay, 2001) and neurological disorders, including mild dementia (Tombaugh, 1996, 1997, but see Teichner & Wagner, 2004). Moreover, age and education have little effect on TOMM scores (Tombaugh, 1996, 1997).

1. Method

1.1. Participants

The Forensic group (age: $M = 32.65$ years, S.D. = 8.46) consisted of 20 men (16/20 minority) from a forensic facility in the southeastern United States. These men were referred to this facility for pre-trial evaluations to assess their competency to stand trial and each had legal proceedings pending. The Psychiatric group (age: $M = 41.00$ years, S.D. = 10.42) consisted of 19 men from an adult inpatient psychiatric unit in the southwestern U.S. and one from another adult inpatient psychiatric unit in the southeastern U.S. In the Psychiatric group, 15/20 were Caucasian. Patients in both groups had similar psychiatric diagnoses on Axis I and Axis II, including, but not limited to, Schizophrenia, Psychosis N.O.S., Bipolar Disorder, and dual diagnoses.

For the Psychiatric group, exclusionary criteria included legal involvement pending (on Axis IV) that suggested secondary gain. For the Forensic group, hospital policy dictated exclusion of individuals with alleged homicide or capital offenses. For either group, individuals who were non-English speaking; had a neurological condition on Axis III; or, who obtained a Global Assessment of Functioning (GAF) Scale score of 30 or less (DSM, Fourth edition, Text revision, American Psychiatric Association, 2000) were excluded.

1.2. Procedure

Each participant was asked to read and sign a release of information and an informed consent statement. Participants were tested individually. The Mini-Mental Status Examination (MMSE; Folstein, Folstein, & McHugh, 1975) was administered first, followed by the M-test (Beaber, Marston, Michelli, & Mills, 1985), a brief test of psychotic malingering. These were followed by administration of Trials 1 and 2 of the TOMM. During the required 15-min
delay between Trial 2 and the Retention Trial, the participants were administered the Wechsler Adult Intelligence Scale–3rd edition (WAIS-III, Wechsler, 1997) Vocabulary and Block Design subtests to obtain an estimated IQ. These were followed by the TOMM Retention Trial and the Antisocial Personality Disorder section of the Structured Clinical Interview for DSM-IV Axis II Screening Questionnaire (SCID-II, First, Gibbon, Spitzer, Williams, & Benjamin, 1997). During testing of the Forensic patients, a security officer was present at all times but remained outside of the room behind a closed door.

2. Results

Eighteen members of the Forensic group had less than a high school education, compared to 13 in the Psychiatric group. Six in the Forensic group met criteria for APD, compared to nine in the Psychiatric group. Seven in the Forensic Group and seven in the Psychiatric group met criteria for malingering on the M-test. None of these differences were significant. Group results on the MMSE, estimated IQ, GAF, and TOMM are presented in Table 1. No group differences emerged on MMSE immediate recall, recall after distraction, GAF; education, number who met criteria for APD and for malingering on the M-test. Significant differences were found for age, MMSE total, and estimated IQ (all d.f. = 38, all p < .05).

A 2 (Forensic, Psychiatric) × 3 (TOMM Trials 1, 2, and Retention) mixed ANOVA (alpha = .05) was used to examine TOMM scores. The Forensic group had significantly lower scores than the Psychiatric group on the three trials ($F_{[1, 38]} = 6.70$, p < .025). The interaction was not significant.

In the Forensic group, significant correlations emerged between total MMSE and TOMM Trial 2 scores ($r = .52$, p < .05) and estimated IQ and TOMM Trial 2 scores ($r = .48$, p < .05), and between total MMSE and TOMM Retention ($r = .48$, p < .05) and estimated IQ and TOMM

Table 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Max. score</th>
<th>Psychiatric group M (S.D.)</th>
<th>Forensic group M (S.D.)</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>32.65 (8.46)</td>
<td>41.00 (10.42)</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>MMSE</td>
<td>Total</td>
<td>30</td>
<td>23.80 (3.98)</td>
<td>26.45 (2.35)</td>
</tr>
<tr>
<td></td>
<td>Immediate</td>
<td>3</td>
<td>3.00 (0.00)</td>
<td>2.90 (0.45)</td>
</tr>
<tr>
<td></td>
<td>Recall after distraction</td>
<td>3</td>
<td>1.00 (1.12)</td>
<td>1.50 (1.05)</td>
</tr>
<tr>
<td>TOMM T1</td>
<td>Est. IQ</td>
<td>74.40 (14.28)</td>
<td>89.85 (19.68)</td>
<td>.007</td>
</tr>
<tr>
<td></td>
<td>GAF</td>
<td>100</td>
<td>45.00 (10.26)</td>
<td>42.50 (11.37)</td>
</tr>
<tr>
<td></td>
<td>TOMM T2</td>
<td>50</td>
<td>39.95 (8.68)</td>
<td>45.10 (3.89)</td>
</tr>
<tr>
<td></td>
<td>TOMM ret.</td>
<td>50</td>
<td>44.05 (9.29)</td>
<td>49.30 (1.66)</td>
</tr>
</tbody>
</table>

* No. of participants scoring below 45: Forensic group = 20; Psychiatric group = 6.
* No. of participants scoring below 45: Forensic group = 6; Psychiatric group = 1.
* No. of participants scoring below 45: Forensic group = 7; Psychiatric group = 1.
* p < .025, Forensic group scores significantly lower than the Psychiatric group on three trials.
Table 2: Frequency of TOMM scores

<table>
<thead>
<tr>
<th>TOMM scores</th>
<th>Forensic Trial 1</th>
<th>Forensic Trial 2</th>
<th>Forensic Trial 3</th>
<th>Psychiatric Trial 1</th>
<th>Psychiatric Trial 2</th>
<th>Psychiatric Trial 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>18–25</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>26–30</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>31–35</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>36–40</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>41–44</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>45–50</td>
<td>6</td>
<td>14</td>
<td>13</td>
<td>14</td>
<td>19</td>
<td>19</td>
</tr>
</tbody>
</table>

Retention (r = .47, p < .05). However, the correlations among these variables in the Psychiatric group were nonsignificant.

The frequency of TOMM scores by trials is presented in Table 2. The difference is striking. Only one member of the Psychiatric group (the same individual) was at or below 45 on Trial 2 and the Retention trial, whereas six members of the Forensic group were at or below 45 on Trial 2, and seven (the same six, plus one other) were at or below this criterion on the Retention Trial.

3. Discussion

This is a differential prevalence design (Rogers, 1997), as opposed to a known groups design, so one would hypothesize that members of the Forensic group would be more likely to have TOMM scores below the recommended cutoffs. The prediction is how likely a member will fall into a group, and not of absolute classification.

As hypothesized, significantly lower TOMM scores were demonstrated by the Forensic group. Also, the frequency of TOMM scores below the recommended cutoff of 45 was clearly reflected in the Forensic group (Trial 2 six versus one, Retention seven versus one) and, as reflected in Table 2, the frequency of low TOMM scores was more prevalent in the Forensic group. This is particularly striking since hospital policy precluded individuals with capital offenses from the Forensic group; consequently, this may be viewed as a rather rigorous test of the TOMM.

The percentage of participants in the Forensic group with TOMM scores below the recommended cutoff on Trial 2 was 30%, and 35% on the Retention trial. These are higher, but consistent with, the base rate of around 22% reported by Mittenberg, Patton, Canyock, and Condit (2002) for criminal cases. Also, in the Psychiatric group, the percentage of participants with TOMM scores below the recommended cutoff was 5%; this is consistent with the adjusted rate of around 8% reported by Mittenberg et al., for medical or psychiatric cases.

One might argue that cognitive status or racial composition, not group membership, was the critical determinant of TOMM scores. However, the participants do not differ on MMSE immediate recall or MMSE recall after distraction. MMSE total and estimated IQ were significantly correlated with TOMM trials in the Forensic group but not in the Psychiatric group. In
a comprehensive review, Green, Lees Haley, and Allen (2002) provided compelling evidence that poor effort reflected in Word Memory Test (WMT) scores was associated with lower scores on other neuropsychological tests. Also, in a comparison of the WMT, the Computerized Assessment of Response Bias and the TOMM in non-head-injury disability claimants (Gervais, Rohling, Green, & Ford, 2004) many more claimants failed the WMT; consequently the TOMM may be viewed as a relatively “easy” symptom validity test.

While previous studies (e.g., Tombaugh, 1996, 1997) showed that the TOMM is largely insensitive to cognitive status, a cautionary note was raised by Teichner and Wagner (2004) in their investigation comparing demented, cognitively impaired, and normal patients on the TOMM. The majority (16/21) of the demented patients fell below the TOMM cutoffs. The authors note that their sample may have been more impaired than Tombaugh’s sample; they also suggested that their criteria for dementia were more rigorous. However, the Forensic group patients in the current study could hardly be viewed as demented.

Limitations to consider in this study include the relatively small sample size, the racial composition of the groups, and the absence of any external assessment of malingering/lack of effort. However, the TOMM identified patients with legal charges pending as more likely to exert poor effort than those without an obvious source of secondary gain.

Acknowledgments

The authors wish to express thanks to James H. Baños and Robert Davis for their critical reading of an earlier version of this manuscript. Thanks also to Beejee Dickson for her help in preparation of this article. This paper is based on the doctoral dissertation of Susan D. Gierok.

References