The Albany Consistency Index for the Test of Memory Malingering†

Jessica H. Gunner1,*, Andrea S. Miele1, Julie K. Lynch2, Robert J. McCaffrey1,2

1Department of Psychology, University at Albany, State University of New York, Albany, NY, USA
2Albany Neuropsychological Associates, Albany, NY, USA

*Corresponding author at: Department of Psychology, University at Albany, State University of New York, Social Sciences 399, 1400 Washington Ave, Albany, NY, 12222, USA.
E-mail address: JGunner@albany.edu (Jessica H. Gunner).

Accepted 10 October 2011

Abstract

The determination of examinee effort is an important component of a neuropsychological evaluation and relies heavily on the use of symptom validity tests (SVTs) such as the Test of Memory Malingering (TOMM) and the Word Memory Test (WMT). Diagnostic utility of SVTs varies. The sensitivity of traditional TOMM criteria to suboptimal effort is low. An index of response consistency across three trials of the TOMM was developed, denoted the Albany Consistency Index (ACI). This index identified a large proportion of examinees classified as optimal effort using traditional TOMM interpretive guidelines but suboptimal effort using the WMT profile analysis. In addition, previous research was extended, demonstrating a relationship between examinee performance on SVTs and neuropsychological tests. Effort classification using the ACI predicted the performance on the Global Memory Index from the Memory Assessment Scales. In conclusion, the ACI was a more sensitive indicator of suboptimal effort than traditional TOMM interpretive guidelines.

Keywords: Symptom validity tests; Test of Memory Malingering; Word Memory Test; Albany Consistency Index; Memory Assessment Scales; Medico-legal; Forensic; Malingering

Introduction

In neuropsychological assessment, examinee performance may range from optimal effort with accurate symptom reporting to suboptimal effort (performance invalidity) with over-reporting of symptoms (Bush et al., 2005). Well-validated measures of effort are critical to the determination of an examinee’s test performance. Failure to perform adequately on effort testing calls into question the validity and reliability of scores from the entire assessment, and conclusions regarding the presence of brain dysfunction cannot be made in the context of failed symptom validity testing (Heilbronner et al., 2009; Hom & Denny, 2002; Iverson & Binder, 2000; Rogers, 1997). Furthermore, Green, Rohling, Lees-Haley, and Allen (2001) found that effort accounted for as much as 53% of the variance in neuropsychological test scores, and that suboptimal effort examinees perform more poorly on neuropsychological tests than those with moderate or severe brain injuries. In fact, researchers have demonstrated that examinees putting forth suboptimal effort perform significantly worse on neuropsychological tests (Beetar & Williams, 1995; Constantinou, Bauer, Ashendorf, & McCaffrey, 2005; van Gorp et al., 1999). For example, examinees with suboptimal effort scored lower on sensory, motor, and attentional measures of the Halstead-Reitan Neuropsychological Battery (HRNB) than those with optimal effort (Mittenberg, Rotholc, Russell, & Heilbronner, 1996). More specifically, effort as determined by the Test of Memory Malingering (TOMM) predicted performance on various neuropsychological measures from the HRNB and Wechsler Adult Intelligence Scale-Revised (WAIS-R) (Constantinou et al., 2005).

† This paper is based upon the first author’s Master thesis. Portions of the paper were presented at the 30th annual meeting of the National Academy of Neuropsychology, Vancouver, BC. Note: For information on the Albany Consistency Index for the Test of Memory Malingering scoring program please contact permission@MHS.com, do not contact the authors.
In addition, suboptimal effort simulators were found to perform significantly lower and had other atypical features to their performance on the Memory Assessment Scales (MAS) than did optimal effort simulators (Beetar & Williams, 1995).

Base rates of suboptimal effort are especially high in forensic and legal work compared with more traditional clinical settings; however, patients seen within a clinical context have also been found to demonstrate suboptimal effort (Larrabee, 2003; Mittenberg, Patton, Canyock, & Condit, 2002). For example, Mittenberg and colleagues (2002) found that 8% of medical cases exhibited suboptimal effort, whereas 29% of personal injury claimants, 30% of disability claimants, and 19% of criminal cases evidenced suboptimal effort. Binder (1993) found similar suboptimal effort rates for litigants, whereas Larrabee (2003) and Mittenberg and colleagues (2002) found that ~40% of mild head injury litigants exhibited probable suboptimal effort. In light of these base rates, any neuropsychological evaluation should be considered incomplete without assessing the examinee’s level of effort (Bush et al., 2005; Iverson & Binder, 2000). Although important, clinical judgment alone is not sufficient to detect suboptimal effort. Tests specifically designed to assess effort are necessary and should be utilized throughout a neuropsychological assessment (van Gorp et al., 1999). Both the National Academy of Neuropsychology (Bush et al., 2005) and the American Academy of Clinical Neuropsychology (Heilbronner et al., 2009) state that all neuropsychological evaluations should examine effort at various points during the assessment, preferably using multiple symptom validity measures from different cognitive domains. Numerous free-standing symptom validity tests (SVTs) have been developed which permit objective identification of suboptimal effort and are not sensitive to neurological deficits. In fact, individuals with brain damage, developmental disabilities, psychiatric problems, and dementia typically perform well on these measures (Heilbronner et al., 2009).

The TOMM is one of the most frequently used SVTs (O’Bryant & Lucas, 2006; Slick, Tan, Strauss, & Hultsch, 2004). A well-researched measure, TOMM cut-off scores have shown to distinguish between suboptimal and optimal effort regardless of age, education, psychological condition, laboratory induced pain, or cognitive impairment (Ashendorf, Constantinoiu, & McCaffrey, 2004; Etherton, Bianchini, Greve, & Ciota, 2005; O’Bryant, Finlay, & O’Jile, 2007; Rees, Tombaugh, & Boulay, 2001; Rees, Tombaugh, Gansler, & Moczynski, 1998; Tombaugh, 1996, 1997). Despite its favor among neuropsychologists, studies have shown that a subset of examinees classified as optimal effort by traditional TOMM criteria is classified as suboptimal effort by Word Memory Test (WMT) profile analysis during the same evaluation (Greve, Binder, & Binachini, 2009; Greve, Ord, Curtis, Bianchini, & Brennan, 2008). For example, Greve and colleagues (2008) found that TOMM cut-off scores had a sensitivity to suboptimal effort of 0.56, whereas WMT profile analysis had a sensitivity of 0.85 in a litigating, mild traumatic brain injury sample. Additionally, Gervais, Rohling, Green, and Ford (2004) reported that while ~10% of claimants were identified as suboptimal effort by TOMM criteria, 30% were similarly identified by the WMT criteria.

Given the TOMM’s popularity among neuropsychologists, clinicians are likely to select this SVT as a measure of effort; however, the traditional TOMM criteria may not be as sensitive to suboptimal effort as criteria from other SVTs such as the WMT. The purpose of the present study was to examine a novel response consistency index to determine whether or not this index might improve the classification accuracy of the TOMM relative to the WMT. First, response consistency was examined across all possible permutations of TOMM trials, and a response consistency index, denoted as the Albany Consistency Index, was developed to identify examinees misclassified as optimal effort (false negatives) using the traditional TOMM criteria. Second, the relationship between effort and neuropsychological test performance was examined using the Global Memory Index of the MAS. This analysis involved a comparison of memory test performance by optimal and suboptimal effort examinees with effort level determined by the standard interpretive TOMM criteria, WMT criteria (profile analysis), and the Albany Consistency Index.

**Method**

**Participants**

After obtaining IRB approval, an archival of analysis was completed on 48 cases from a private neuropsychological practice that were conducted for medico-legal reasons (i.e., compensation-seeking, litigation, or disability claims). This was a mixed neurological sample, although all but five examinees had a history of mild traumatic brain injury and met the mild traumatic brain injury criteria from the American Congress of Rehabilitation Medicine (ACRM) Mild Traumatic Brain Injury Committee of the Head Injury Interdisciplinary Special Interest Group (Committee on Mild Traumatic Brain Injury, 1993). Five examinees presented with other medical conditions, diagnosed by their treating neurologist. The diagnoses were fibromyalgia (n = 2); transient ischemic attack (n = 1); brain stem stroke and major depressive disorder (n = 1); and multiple sclerosis (n = 1).
**Materials**

The archival data consisted of scores from the HRNB (Reitan and Wolfson, 1993) and the MAS (Williams, 1991), a memory test that contains four summary index scores: Short-term Memory, Verbal Memory, Visual Memory, and Global Memory. Symptom validity measures consisted of the TOMM (Tombaugh, 1996) and WMT (Green, 2003).

**Procedure: Creating the Albany Consistency Index**

All examinees completed both the WMT and the TOMM. Examinee data were classified into eight groups for subsequent analyses. Groups 7 and 8 were classified using the Albany Consistency Index, which is described below.

- Group 1: TOMM—Optimal Effort
- Group 2: TOMM—Suboptimal Effort
- Group 3: WMT—Optimal Effort
- Group 4: WMT—Suboptimal Effort
- Group 5: TOMM—Optimal Effort; WMT—Optimal Effort
- Group 6: TOMM—Optimal effort; WMT—Suboptimal Effort
- Group 7: Albany Consistency Index—Optimal Effort
- Group 8: Albany Consistency Index—Suboptimal Effort

TOMM individual item responses were used to create an Index of consistently incorrect, inconsistent, or consistently correct responses in an effort to quantify the consistency of examinees’ responses across Trial 1, Trial 2, and the Retention Trial. Responses for items during each of these trials were coded as either one (1) for a correct response or zero (0) for an incorrect response. Coded responses were then evaluated across four different groupings of the three TOMM trials: Trial 1 and Trial 2 (Index A), Trial 1 and the Retention Trial (Index B), Trial 2 and the Retention Trial (Index C), and across all three trials (Index D). For each index, examinees’ responses were then identified as consistently incorrect, inconsistent, or consistently correct for each of the 50-items. Consistently incorrect responses were those answered incorrectly on both trials of Indices A, B, or C, or all three trials of Index D. Inconsistent responding was defined as one correct response and one incorrect response in any order on Indices A, B, and C. Inconsistent responding across three trials for Index D was defined as either responding to the item correctly on one trial and incorrectly on the other two trials, or correctly on two trials and incorrectly on one trial. Consistently correct responding was defined as responding correctly across both trials for Indices A, B, or C, or across all three trials of Index D. A description of each of the four indices is presented in Table 1.

The response consistency patterns of examinees in Groups 5 and 6 were examined to assess for any variation between (i) examinees who exhibited optimal effort on both the TOMM and WMT criteria and (ii) examinees whose effort was optimal based on the TOMM criteria but whose effort was suboptimal based on WMT criteria. In this study, 16 examinees exhibited optimal effort based on TOMM criteria but suboptimal effort based on WMT profile analysis. These “optimal effort” performances identified using the TOMM were considered to represent false negatives; that is, the TOMM criteria failed to detect suboptimal effort that was identified by the WMT criteria. Twenty-three examinees exhibited optimal effort based on both the TOMM and WMT criteria. The TOMM performances in this group were considered true negatives. The frequency of examinees in Groups 5 and 6 responding consistently incorrect, inconsistent, and consistently correct was then compared for each index. After examining the frequency graphs for false positives and false negatives for all groups across all indices, Index D was selected for development of a cutoff score, and from this point forward will be referred to as the Albany Consistency Index. This index was selected because none of the examinees exhibiting optimal effort based on the TOMM...
on the TOMM and WMT criteria responded inconsistently on ten or more items of the Albany Consistency Index. As such, the cutoff score for the identification of suboptimal effort using the Albany Consistency Index was set at ≥10 inconsistent responses. Examinees who responded inconsistently on ≥10 items on the Albany Consistency Index were classified as putting forth suboptimal effort (Group 8), while those who responded inconsistently on <10 items on the Albany Consistency Index were classified as putting forth optimal effort (Group 7).

Procedure: Remaining Analyses

Statistical tests included three logistic regressions and univariate ANOVAs. Logistic regression was used to examine the accuracy with which the standard TOMM criteria and the Albany Consistency Index classified optimal and suboptimal effort relative to WMT criteria. This was used to determine the predictive validity of the Albany Consistency Index over and above the standard TOMM criteria in determining suboptimal effort as established by the WMT criteria. Univariate tests were utilized to examine the differences on the Global Memory Index of the MAS between optimal and suboptimal effort groups. Three separate univariate tests were conducted to compare the MAS test performances between the effort groups as determined by the standard TOMM criteria, Albany Consistency Index, and WMT criteria.

Results

Demographics and Classification Accuracy

Detailed demographic characteristics for all groups can be found in Table 2. Examinees’ average age and years of education were 45.3 (9.8) and 14.2 (2.9), respectively. Of these subjects, 23 were male and 45 were right handed. The average scores on the WAIS-R Verbal IQ (VIQ), Performance IQ (PIQ), and Full-scale IQ (FSIQ) scores were 97 (12), 94 (11), and 95 (11), respectively. Chi-square analyses showed no significant differences in age, education, gender, handedness, or WAIS-R scores between optimal and suboptimal effort examinees based on standard TOMM criteria, Albany Consistency Index, or WMT criteria.

One examinee was classified as suboptimal effort on the TOMM but classified as optimal effort on the WMT (false positive) and 16 examinees were classified as optimal effort on the TOMM but suboptimal effort on the WMT (false negatives). This indicated that, based on the WMT criteria, the standard TOMM criteria had a sensitivity of 0.33 and a specificity of 0.96 (see Fig. 1). The Albany Consistency Index correctly identified suboptimal effort in nine of the 16 examinees (56%) classified as

Table 2. Means, Standard Deviations, and ranges for the eight participant groups

<table>
<thead>
<tr>
<th>Group</th>
<th>SVT</th>
<th>TOMM Pass (n = 48)</th>
<th>TOMM Fail (n = 9)</th>
<th>WMT Pass (n = 48)</th>
<th>WMT Fail (n = 24)</th>
<th>TOMM Pass and WMT Pass (n = 23)</th>
<th>TOMM Pass and WMT Fail (n = 16)</th>
<th>Pass Consistency Index (n = 30)</th>
<th>Fail Consistency Index (n = 18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>M (SD)</td>
<td>45 (10)</td>
<td>46 (10)</td>
<td>41 (10)</td>
<td>45 (10)</td>
<td>44 (10)</td>
<td>46 (9)</td>
<td>44 (10)</td>
<td>49 (7)</td>
</tr>
<tr>
<td>Gender (M/F)</td>
<td>23/25</td>
<td>21/18</td>
<td>2/7</td>
<td>23/25</td>
<td>14/10</td>
<td>9/15</td>
<td>13/10</td>
<td>8/8</td>
<td>14/16</td>
</tr>
<tr>
<td>Education M (SD)</td>
<td>14.2 (2.9)</td>
<td>14.3 (2.8)</td>
<td>14.0 (3.3)</td>
<td>14.2 (2.9)</td>
<td>14.5 (3.0)</td>
<td>14.0 (2.8)</td>
<td>14.5 (3.1)</td>
<td>13.9 (2.5)</td>
<td>14.6 (2.9)</td>
</tr>
<tr>
<td>Handedness (R/L)</td>
<td>45/5</td>
<td>36/3</td>
<td>9/0</td>
<td>45/3</td>
<td>23/1</td>
<td>22/2</td>
<td>22/1</td>
<td>14/2</td>
<td>28/2</td>
</tr>
<tr>
<td>WAIS-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIQ M (SD)</td>
<td>97 (12)</td>
<td>98 (12)</td>
<td>91 (12)</td>
<td>97 (12)</td>
<td>101 (11)</td>
<td>93 (12)</td>
<td>101 (11)</td>
<td>94 (12)</td>
<td>100 (12)</td>
</tr>
<tr>
<td>PIQ M (SD)</td>
<td>94 (11)</td>
<td>96 (11)</td>
<td>86 (10)</td>
<td>94 (11)</td>
<td>98 (12)</td>
<td>90 (10)</td>
<td>99 (12)</td>
<td>91 (9)</td>
<td>98 (11)</td>
</tr>
<tr>
<td>FSIQ M (SD)</td>
<td>95 (11)</td>
<td>97 (11)</td>
<td>88 (11)</td>
<td>95 (11)</td>
<td>99 (11)</td>
<td>91 (10)</td>
<td>100 (10)</td>
<td>93 (10)</td>
<td>98 (11)</td>
</tr>
</tbody>
</table>
optimal effort using the TOMM criteria but as suboptimal effort based on WMT criteria. In addition, while identifying 56% of the false negatives, the Albany Consistency Index did not result in any false positives. In fact, this index correctly classified the one false positive from the standard TOMM criteria (i.e., the examinee was correctly classified as optimal effort). Thus, evaluating effort using the Albany Consistency Index increased the sensitivity of the TOMM cut-off scores from 0.31 to 0.71 and also increased specificity to 1.00 (see Fig. 2).

**Validation of the Albany Consistency Index**

Logistic regression was used to evaluate the incremental validity of the Albany Consistency Index. Three direct logistic regressions were executed. The first (Model 1) utilized only the standard TOMM criteria as an independent variable in order to classify examinees into optimal and suboptimal effort groups as defined by the WMT criteria. Table 3 shows the regression coefficients, Wald statistics, odds ratios, and 95% confidence intervals for the odds ratios for the TOMM criteria as a sole predictor variable. This model had a $\chi^2$-value of 59.08. According to the Wald criterion, the TOMM criteria reliably predicted effort as determined by the WMT criteria, $\chi^2 (1, N = 48) = 4.85, p = .028$. The variance in effort accounted for by the standard TOMM criteria was small (McFadden’s $D = 0.11$; 95% confidence interval of the effect = 2.44). The overall accuracy rate was 64.6%, in which 95.8% of the optimal effort examinees and only 33.3% of the suboptimal effort examinees were correctly classified. This model produced an odds ratio of 11.50, which indicates that a failing performance on the WMT is 11.5 times more likely if an examinee has failed the TOMM.

![Fig. 1. Calculating classification accuracy for the TOMM based on WMT classification.](image1)

<table>
<thead>
<tr>
<th>TOMM</th>
<th>Pass (Optimal effort)</th>
<th>Fail (Suboptimal effort)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMT</td>
<td>True negative</td>
<td>False positive</td>
</tr>
<tr>
<td></td>
<td>23 examinees</td>
<td>1 examinees</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Albany Consistency Index</th>
<th>Pass (Optimal effort)</th>
<th>Fail (Suboptimal effort)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMT</td>
<td>True negative</td>
<td>False positive</td>
</tr>
<tr>
<td></td>
<td>24 examinees</td>
<td>0 examinees</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Albany Consistency Index</th>
<th>Pass (Optimal effort)</th>
<th>Fail (Suboptimal effort)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMT</td>
<td>False negative</td>
<td>True positive</td>
</tr>
<tr>
<td></td>
<td>7 examinees</td>
<td>17 examinees</td>
</tr>
</tbody>
</table>

![Fig. 2. Calculating classification accuracy for the Albany Consistency Index based on WMT classification.](image2)
The second direct logistic regression (Model 2) examined the classification accuracy of the standard TOMM criteria and the Albany Consistency Index together to determine whether the Albany Consistency Index increased the correct classification of suboptimal and optimal effort based on the WMT criteria compared with the standard TOMM criteria alone (Model 1). Table 4 shows the results from the logistic regression for both the standard TOMM criteria and the Albany Consistency Index. According to the Wald criterion, the Albany Consistency Index reliably predicted effort as determined by the WMT, \(\chi^2 (1, n = 48) = 7.38, p = .007\), while the standard TOMM criteria did not significantly predict effort, \(\chi^2 (1, n = 48) = 0.00, p = 1.000\). The variance in effort accounted for by the standard TOMM criteria and Albany Consistency Index in this second model was larger than that obtained in the previous logistic regression (McFadden’s \(D = 0.29\); 95% confidence interval for the effect = 3.09). This analysis correctly classified 91.7% of the optimal effort examinees and 66.7% of the suboptimal effort examinees, with an overall accuracy rate of 79.2% compared with 64.6% in the first model. The odds ratio of 1.00 associated with the TOMM criteria indicates that suboptimal effort as determined by these criteria did not increase the likelihood of also having suboptimal effort as determined by the WMT criteria. However, according to the second odds ratio, identifying suboptimal effort on the WMT is 22.00 times more likely if an examinee exhibited suboptimal effort on the Albany Consistency Index. The second model had an overall \(\chi^2\)-value of 47.35, which is significantly smaller than the first model with a \(\chi^2\) value of 59.08 (\(p < .05\)). Statistically, the use of both the standard TOMM criteria and Albany Consistency Index as predictors of WMT performance is more accurate in classifying effort than the TOMM alone.

The final logistic regression (Model 3) analyzed the utility of the Albany Consistency Index alone in predicting effort as determined by the WMT criteria. See Table 5 for results from the logistic regression including only the Albany Consistency Index. According to the Wald criterion, the Albany Consistency Index reliably predicted effort as determined by the WMT criteria, \(\chi^2 (1, N = 48) = 13.04, p < .005\). The variance in effort accounted for by the Albany Consistency Index was identical to the second model that utilized both the standard TOMM and the Albany Consistency Index (McFadden’s \(D = 0.29\); 95% confidence interval for the effect = 3.09). Classification results were also identical to the second model with 91.7% of the optimal effort examinees and 66.7% of the suboptimal effort examinees correctly classified, for an overall accuracy rate of 79.2%. Similar to the second model, the odds ratio of 22.00 indicates that suboptimal effort based on the WMT criteria is 22 times more likely if an examinee exhibited suboptimal effort on the Albany Consistency Index. The second model had an overall \(\chi^2\)-value of 47.35, which is significantly smaller than the first model with a \(\chi^2\)-value of 59.08 (\(p < .05\)). Statistically, the use of both the standard TOMM criteria and Albany Consistency Index as predictors of WMT performance is more accurate in classifying effort than the TOMM alone.

### Relationship of Effort to Memory Performance

Three univariate ANOVAs were conducted to compare the performance on the Global Memory Index of the MAS between examinees with optimal versus suboptimal performance on the TOMM, Albany Consistency Index, and WMT. The Short-Term Memory Index, Verbal Memory Index, and Visual Memory Index were all significantly correlated with the Global Memory Index (\(r = .46, p < .0005\), \(r = .88, p < .0005\), and \(r = .86, p < .0005\), respectively); thus, the Short-term Memory, Verbal Memory, and Visual Memory Indices were regressed on the Global Memory Index. This resulted in a significant adjusted \(R^2\) value of .995 (\(F (3, 48) = 3,695.3, p < .0005\), indicating that these three indices accounted for 99.5% of the

<table>
<thead>
<tr>
<th>Variable</th>
<th>Wald chi-square</th>
<th>Odds ratio</th>
<th>95% Confidence interval for odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOMM</td>
<td>2.442</td>
<td>11.500</td>
<td>1.307 101.182</td>
</tr>
<tr>
<td>Albany Consistency Index</td>
<td>3.091</td>
<td>22.000</td>
<td>2.364 204.759</td>
</tr>
</tbody>
</table>

Table 3. Logistic regression analysis of effort determined by the WMT as a function of suboptimal and optimal effort based on the standard TOMM criteria

<table>
<thead>
<tr>
<th>Variable</th>
<th>Wald chi-square</th>
<th>Odds ratio</th>
<th>95% Confidence interval for odds ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOMM</td>
<td>2.442</td>
<td>11.500</td>
<td>1.307 101.182</td>
</tr>
<tr>
<td>Albany Consistency Index</td>
<td>3.091</td>
<td>22.000</td>
<td>2.364 204.759</td>
</tr>
</tbody>
</table>

Table 4. Logistic regression analysis of effort determined by the WMT as a function of suboptimal and optimal effort based on the standard TOMM criteria and the Albany Consistency Index
variance in the Global Memory Index scores. This was expected, as the Global Memory Index is comprised of scores from the other three indices. Based on this analysis, only the Global Memory Index of the MAS was used for group comparisons to avoid redundancy and multicollinearity.

Group means, ranges, and standard deviations on the Global Memory Index can be found in Table 6. ANOVAs were used instead of t-tests to obtain the eta-squared values (a measure of effect size) for each comparison. Eta-squared values provide the variance in the dependent variable (Global Memory Index) uniquely attributable to the predictor variable (TOMM, Albany Consistency Index, or WMT). Following a Bonferroni correction, the alpha level was adjusted to $p < .0166$. The results indicated that there were no significant differences in scores on the Global Memory Index between examinees with optimal versus suboptimal effort as determined by the standard TOMM criteria, $F(1, 47) = 5.06, p = .029$. Effort based on the standard TOMM criteria accounted for only 9.7% of the variance in the Global Memory Index (eta-squared obtained $= .097$).

In contrast, there were significant differences on Global Memory Index scores between optimal and suboptimal effort examinees as determined by the Albany Consistency Index, $F(1, 47) = 7.95, p = .007$. Examinees exhibiting optimal effort on the Albany Consistency Index performed significantly better on the Global Memory Index compared with those putting forth suboptimal effort. In this analysis, effort as determined by the Albany Consistency Index accounted for 14.5% of the variance in scores on the Global Memory Index (eta-squared obtained $= .145$).

There were also significant differences in scores on the Global Memory Index between optimal versus suboptimal effort examinees as determined by WMT criteria, $F(41, 5) = 16.33, p < .005$, with optimal effort examinees performing significantly better on the Global Memory Index. Effort as indexed by WMT criteria accounted for 26.6% of the variance in scores on the Global Memory Index (eta-squared obtained $= .266$).

Table 6. Means, standard deviations, and ranges for the Global Memory Index of the Memory Assessment Scale

<table>
<thead>
<tr>
<th>Variable</th>
<th>TOMM optimal effort</th>
<th>TOMM suboptimal effort</th>
<th>Albany Consistency Index optimal effort</th>
<th>Albany Consistency Index suboptimal effort</th>
<th>WMT optimal effort</th>
<th>WMT suboptimal effort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>87.95</td>
<td>75.00</td>
<td>90.09</td>
<td>77.82</td>
<td>93.79</td>
<td>77.65</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>14.74</td>
<td>15.78</td>
<td>15.13</td>
<td>13.19</td>
<td>14.30</td>
<td>13.02</td>
</tr>
</tbody>
</table>

Discussion

This study examined whether a response consistency index would improve the TOMM’s sensitivity to suboptimal effort based on WMT classifications. The Albany Consistency Index identified 56% of the examinees who exhibited optimal effort on the TOMM but suboptimal effort on the WMT, while not identifying any optimal effort examinees as putting forth suboptimal effort. The Albany Consistency Index increased the sensitivity of the traditional TOMM cut-off scores from 0.31 to 0.71 while maintaining high specificity (0.96–1.00). These results indicate that when effort was determined by the WMT criteria, the Albany Consistency Index provided more accurate classification of suboptimal effort and perfect classification of optimal effort compared with the traditional TOMM criteria.

Three logistic regressions were conducted to determine the ability of the standard TOMM criteria alone (Model 1), standard TOMM criteria and Albany Consistency Index together (Model 2), and the Albany Consistency Index alone (Model 3) to predict effort as classified by the WMT criteria. Model 1 indicated that the standard TOMM criteria are much better at identifying optimal rather than suboptimal effort. Model 2 showed an increased correct classification of effort groups and also indicated that the standard TOMM criteria did not significantly increase the predictive accuracy of the Albany Consistency Index. Thus, Model 3, which used only the Albany Consistency Index, was the simplest and most accurate model for classifying effort as determined by the WMT criteria. This model resulted in the exact same $X^2$ value, classification results, and variance accounted for in the WMT as obtained in Model 2, indicating that the standard TOMM criteria did not increase the incremental...
validity of the Albany Consistency Index. Additionally, those examinees identified as suboptimal effort by the Albany Consistency Index were almost two times more likely than those identified as suboptimal effort by the standard TOMM criteria to have also been identified as suboptimal effort by the WMT criteria. Thus, together, logistic regression and classification accuracy statistics demonstrate the utility of the Albany Consistency Index in identifying suboptimal effort.

A second goal of this study was to replicate and expand prior research, showing that suboptimal effort examinees perform lower than the optimal effort examinees on neuropsychological tests (Beetar & Williams, 1995; Constantinou et al., 2005). The results of the current study demonstrated that suboptimal effort examinees performed significantly lower on the Global Memory Index of the MAS when using either the WMT criteria or the Albany Consistency Index to classify effort. However, when classifying effort using the standard TOMM criteria, suboptimal and optimal effort groups obtained similar Global Memory Index scores. It is likely that effort as determined by these criteria were not predictive of scores on the Global Memory Index due to the TOMM criteria’s limited ability to accurately classify suboptimal effort. Examinees identified as optimal rather than suboptimal effort by the WMT criteria or the Albany Consistency Index were more likely to have higher scores on the Global Memory Index. Therefore, effort classification by both WMT criteria and the Albany Consistency Index was predictive of scores on the Global Memory Index. Effort as determined by the WMT criteria, the Albany Consistency Index, and standard TOMM criteria accounted for 26.6%, 14.5%, and only 9.7% of the variance on the Global Memory Index, respectively.

Limitations and Future Directions

Results from the logistic regressions and univariate ANOVAs may only generalize to a similar litigating sample. In addition, when utilizing logistic regression, it is important to consider both correlations between the predictor variables as well as between the predictor and outcome variables. Although within acceptable limits for a logistic regression, the variables used in the three models (the TOMM, Albany Consistency Index, and WMT) were moderately correlated and, therefore, results from the inferential statistics may overestimate their classification ability. For this reason, causal relationships should not be inferred from the above results. Additionally, due to conducting multiple comparisons, a Bonferroni correction was utilized to protect against Type I error; however, these corrected alpha values may be overly stringent and mask true differences between the groups. Finally, although computing the consistency scores from individual TOMM data may require additional time, a computer spreadsheet application can provide assistance. According to Bush and colleagues (2005), increased time allocated to the accurate assessment of effort is both necessary and justified.

Future research may benefit from utilizing a larger sample that would allow for more powerful statistical tests. In addition, the Albany Consistency Index should be cross-validated using both litigating and clinical samples.

Conclusions

The major findings of this study validate the use of the Albany Consistency Index for the TOMM and support previous research, showing that suboptimal effort examinees perform more poorly on neuropsychological tests than do optimal effort examinees (Beetar & Williams, 1995; Constantinou et al., 2005). Examinees performing better on the Global Memory Index were significantly more likely to have optimal effort scores on both the Albany Consistency Index and WMT compared with those who performed poorly on the Global Memory Index. More importantly, the Albany Consistency Index identified a large proportion of examinees who exhibited optimal effort based on the traditional TOMM criteria but suboptimal effort based on the WMT criteria. The Albany Consistency Index more than doubled the sensitivity of the traditional TOMM criteria for identifying suboptimal effort, thereby demonstrating that it is a more sensitive indicator of effort than the traditional TOMM criteria.

Conflict of Interest

The first author would like to acknowledge the generous travel grant support from the Neuropsychology Division of the New York State Psychological Association.

Acknowledgements

The authors are grateful to Cecil R. Reynolds, former editor of Archives of Clinical Neuropsychology and current editor of Psychological Assessment, who served as the guest action editor. This manuscript was submitted to the blind peer review
References


