



Research Article

What are the predictors of TOMM failure in clinical TBI populations? A retrospective analysis

Thomas M. Gates^{1,2} , Kasey Metcalf¹, Mark Sabaz³, Kate Young¹, Thushara Anandakumar^{1,3} and Grahame K. Simpson^{1,2,4}

¹Liverpool Brain Injury Rehabilitation Unit, Liverpool Hospital, Sydney, Australia, ²Brain Injury Rehabilitation Research Group, Ingham Institute for Applied Medical Research, Sydney, Australia, ³Department of Psychology, Sydney Children's Hospital, Sydney, Australia and ⁴John Walsh Centre for Rehabilitation Research, Faculty of Medicine and Health, University of Sydney, Sydney, Australia

Abstract

Objectives: To determine base rates of invalid performance on the Test of Memory Malingering (TOMM) in patients with traumatic brain injury (TBI) undertaking rehabilitation who were referred for clinical assessment, and the factors contributing to TOMM failure. **Methods:** Retrospective file review of consecutive TBI referrals for neuropsychological assessment over seven years. TOMM failure was conventionally defined as performance <45/50 on Trial 2 or Retention Trial. Demographic, injury, financial compensation, occupational, and medical variables were collected. **Results:** Four hundred and ninety one TBI cases (Median age = 40 years [IQR = 26–52], 79% male, 82% severe TBI) were identified. Overall, 48 cases (9.78%) failed the TOMM. Logistic regression analyses revealed that use of an interpreter during the assessment (adjusted odds ratio [aOR] = 8.25, 95%CI = 3.96–17.18), outpatient setting (aOR = 4.80, 95%CI = 1.87–12.31) and post-injury psychological distress (aOR = 2.77, 95%CI = 1.35–5.70) were significant multivariate predictors of TOMM failure. The TOMM failure rate for interpreter cases was 49% (21/43) in the outpatient setting vs. 7% (2/30) in the inpatient setting. By comparison, 9% (21/230) of non-interpreter outpatient cases failed the TOMM vs. 2% (4/188) of inpatient cases. **Conclusions:** TOMM failure very rarely occurs in clinical assessment of TBI patients in the inpatient rehabilitation setting. It is more common in the outpatient setting, particularly in non-English-speaking people requiring an interpreter. The findings reinforce the importance of routinely administering stand-alone performance validity tests in assessments of clinical TBI populations, particularly in outpatient settings, to ensure that neuropsychological test results can be interpreted with a high degree of confidence.

Key words: Cross-cultural comparison; Memory; Neuropsychological assessment; Performance validity; Rehabilitation; Traumatic brain injury

(Received 11 October 2021; final revision 21 April 2022; accepted 3 May 2022; First Published online 11 July 2022)

Performance validity tests (PVTs) are used by clinical neuropsychologists to verify that examinees are sufficiently motivated and that their performance on neuropsychological testing genuinely reflects their true cognitive capacity. This is necessary to ensure that conclusions and recommendations drawn from test results are valid. Research into performance validity assessment has traditionally focused on medico-legal, forensic, and veteran settings, with estimates of non-credible performance in these populations typically ranging from 30–50% (e.g., Green et al., 2001; Larrabee et al., 2009). However, the past 10 years has seen a shift towards PVT research occurring increasingly in clinical settings (Suchy, 2019).

The recent update to the 2009 American Academy of Clinical Neuropsychology (AACN) Consensus Statement on Validity Assessment (Sweet, Heilbronner, et al., 2021) recognized that invalid responding occurs in persons presenting for routine medical care in non-forensic and non-litigating settings, even in the absence of clearly defined internal and external incentives. Attitudes towards PVTs have become increasingly positive among

US neuropsychologists over the past five years, moving towards near universal acceptance of the notion that using stand-alone and embedded PVTs in practice is well-supported by the available empirical evidence (SweetKlipfel, et al., 2021). There is also strong agreement that clinical judgment is often inaccurate and therefore insufficient as a means of detecting invalid responding by itself (Sweet, Heilbronner, et al., 2021). However, these sentiments have not been uniformly translated into clinical practice worldwide, with one recent example provided by a survey of Australian neuropsychologists indicating PVTs were employed in less than half of all clinical assessments (Uiterwijk et al., 2021).

For these reasons it is essential to continue building the evidence base underpinning PVT use in clinical practice. To date, individual clinical studies (e.g., Johnson-Greene et al., 2013; Kemp et al., 2008; Sherer et al., 2020; Wodushek & Domen, 2020) as well as a recent systematic review (McWhirter et al., 2020) have found significant variations in base rates of PVT failure across different clinical populations, ranging from less than 20% to higher than 25%. Moreover, higher rates (more in line with

Corresponding author: Thomas M. Gates, email: thomas.gates@health.nsw.gov.au

Cite this article: Gates T.M., Metcalf K., Sabaz M., Young K., Anandakumar T., & Simpson G.K. (2023) What are the predictors of TOMM failure in clinical TBI populations? A retrospective analysis. *Journal of the International Neuropsychological Society*, 29: 336–345, <https://doi.org/10.1017/S1355617722000339>

medico-legal settings) have been reported when external incentives were present (Fox, 2011; Johnson-Greene et al., 2013; Martin & Schroeder, 2020; Sullivan et al., 2007). Factors associated with these variable base rates in clinical populations include greater heterogeneity in the nature and severity of the presenting medical condition(s); grouping of disparate clinical groups into a single cohort; the context of the assessment/referral question; the types of external incentives present and the way motivating factors evolve over time; as well as differences in the specific PVTs selected and cut-off thresholds used (Martin & Schroeder, 2020; Schroeder et al., 2019).

To address these issues, more studies focusing on specific medical conditions are needed. Traumatic brain injury (TBI) is one of the most commonly referred conditions for neuropsychological assessment (Sweet, Klipfel, et al., 2021). While TBI has been reasonably well represented within the clinical PVT literature to date, studies have tended to focus on the milder end of the injury spectrum. Much less is known about the base rates of PVT failure in those with more severe injuries.

The current study employed the Test of Memory Malingering (TOMM; Tombaugh, 1996), which is the most commonly used stand-alone PVT (Martin et al., 2015). The TOMM has been extensively validated in many adult and child populations (for an in-depth review, see Martin et al., 2020). It is widely recognized as being insensitive to moderate/severe TBI (Rees, Tombaugh, & Boulay, 2001), with the Martin et al. (2020) meta-analysis reporting a weighted mean specificity of .91 using the traditional cut-off threshold of scores below 45/50 on Trial 2 or the Retention Trial (Bashem et al., 2014; Greve et al., 2006; Rees et al., 1998).

More studies are needed that prospectively and systematically administer PVTs to all patients regardless of setting (e.g., inpatient or outpatient). To our knowledge, only one study (Moore & Donders, 2004) has systematically examined TOMM performance within the clinical setting of a brain injury rehabilitation service. In that study, the rate of TOMM failure was 8% (11/132). Prior psychiatric history and financial compensation-seeking were associated with nearly four-fold increased odds of PVT failure. To extend these results, it would be ideal to examine a cohort with less variation in terms of injury severity, as Moore and Donders' (2004) sample comprised 55% mild TBI cases and 45% moderate/severe TBI. In addition, comprehensive data on TOMM performance in TBI inpatients is lacking, as attested by the fact that around 90% of Moore and Donders' (2004) sample were assessed as outpatients. Clinical studies of inpatients have mostly been conducted in psychiatric samples and have generally found lower rates of PVT failure relative to outpatient settings (e.g., Gierok et al., 2005; Lee, et al., 2021; Rees et al., 2001). Further study in TBI samples with adequate representation of inpatient as well as outpatient settings is therefore desirable.

Finally, both the 2021 update to the 2009 AACN Consensus Statement (Sweet, Heilbronner, et al., 2021) and revised Slick et al. criteria for malingered cognitive dysfunction (Sherman et al., 2020) highlight the need for ongoing validation of performance validity tests in patient populations from diverse language, culture, education, and socio-economic backgrounds (also see Nijdam-Jones & Rosenfeld, 2017; Salazar et al., 2007; Strutt et al., 2012). The current study was able to explore the influence of socio-cultural and language factors on TOMM performance in a severely injured TBI sample. Liverpool Hospital (the study setting) services the highly culturally diverse South-Western region of Sydney, Australia. According to the most recent Australian census in 2016, 53% of people living in this region were born overseas

(vs. 33% of the Australian population) and 64% of households spoke a language other than English at home (vs. 22%) (Australian Bureau of Statistics, 2017).

In summary, the aims of the current study were twofold: (1) to determine the base rates of TOMM failure in clinical TBI cases consecutively referred for neuropsychological assessment within an adult brain injury rehabilitation service; and (2) to identify potential injury-related and other medical, service setting, and socio-demographic/cultural factors associated with TOMM failure in this population.

Methods

Participants

The current study was set within the Brain Injury Rehabilitation Unit (BIRU) of Liverpool Hospital, Sydney, Australia. The BIRU provides long-stay inpatient and community-based (outpatient) rehabilitation to patients with severe TBI, with clinical neuropsychologists providing comprehensive assessments in both settings. Neuropsychological assessment results were obtained through a consecutive file review of all inpatient and outpatient TBI referrals made to BIRU neuropsychologists between July 2012 and December 2019. During this period, a policy was adopted within the BIRU whereby the TOMM was to be administered in all neuropsychological assessments bar a few exceptional circumstances (more detail is provided in Figure 1).

The study was conducted in accordance with the Declaration of Helsinki. Ethical approval was granted by the South-Western Sydney Local Health District Human Research Ethics Committee. Given the retrospective nature of the study it was approved as a low/negligible risk project, and the requirement for informed consent from cases was waived.

Initially 576 TBI cases were identified, of which 85 were excluded through two stages of screening, resulting in a final sample of 491 cases. Of the excluded cases, 21 were due to TOMM administration being considered not appropriate given the circumstances of the referral and/or severe cognitive deficits precluding valid administration, while for an additional 64 cases TOMM administration was considered to have been appropriate, but it was not administered (deviating from the agreed-upon policy) or invalidated or abandoned due to a variety of clinical factors that arose during the assessment (see Figure 1 for more details). The excluded cases were compared to the final sample in relation to the main predictive variables to ensure that the screening process did not lead to significant selection biases that could threaten the internal validity of the study.

Materials

Test of Memory Malingering (TOMM; Tombaugh, 1996)

The TOMM is a visual recognition memory test designed to help clinicians distinguish feigned memory impairment from genuine. It consists of two learning trials and an optional retention trial. During the learning trials the examinee is presented with 50 common objects sequentially. They are then required to make a forced choice between a target (one of the original 50 items) or a foil, selecting the item they believe they had seen previously. The Retention Trial occurs 15 min after completion of Trial 2, consisting of the forced-choice component without additional presentation of the original 50 items.

For clinical purposes, administration of the TOMM was modified such that if a case passed Trial 1 (i.e., scoring $\geq 45/50$)

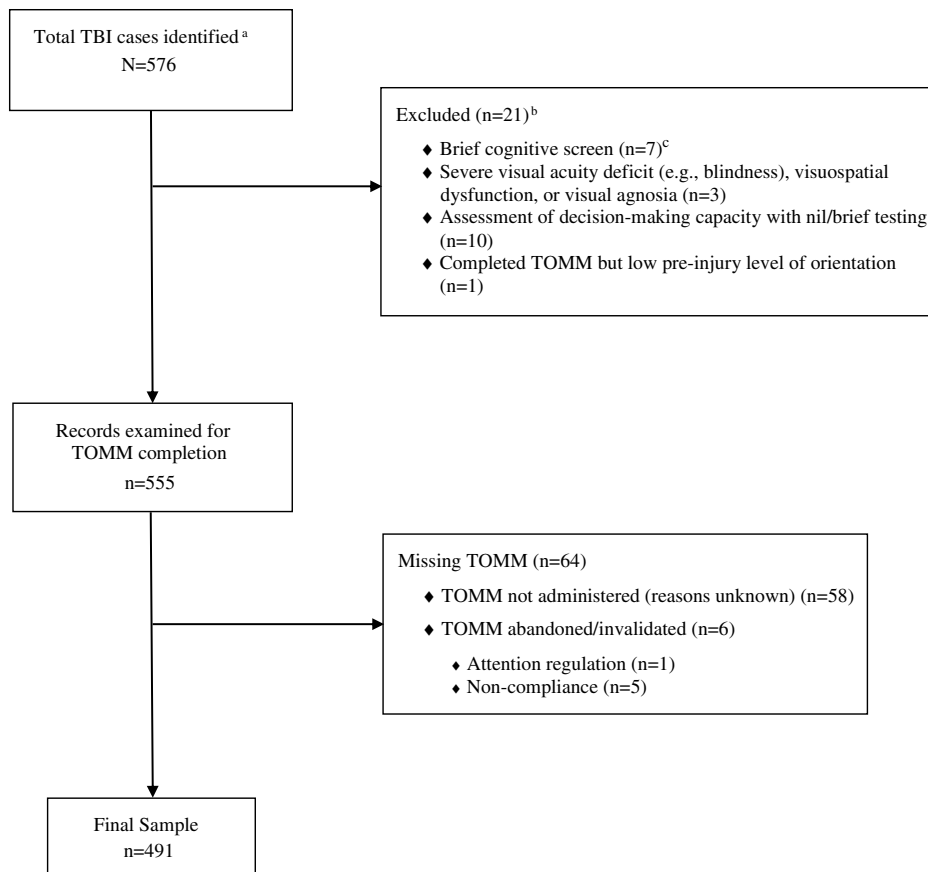


Figure 1. Neuropsychological assessment record screening process. ^aFor cases that completed multiple assessments over study period, only one assessment record was considered this study. This was either the first assessment in which TOMM was validly administered, or the first completed assessment if the TOMM was not administered during any assessments. ^bCases were excluded at this stage of screening if TOMM administration was considered to have been not appropriate (either in the context of the referral, or if the nature and severity of the case's cognitive deficits precluded valid administration). ^cCases were excluded if the only cognitive testing administered was one of the following screening measures: MoCA, MMSE, ACE-R/ACE-III, or FAB. Cases were not excluded at this stage if a more comprehensive cognitive screening battery was completed (e.g., RBANS), or if one or more of these measures was completed alongside other standard neuropsychological tests.

then subsequent trials were not administered. All of these cases were coded as having passed the TOMM. Similarly, if cases were administered and passed Trial 2, the Retention Trial was not administered, and these cases were also coded as having passed the TOMM. Overall, Trial 1 was administered to all 491 cases, Trial 2 to 194 (39%) cases, and the Retention Trial to 68 (14%) cases.

TOMM failure was the primary study outcome. It was defined as a score of <45/50 on either Trial 2 or the Retention Trial, following the conventional cut-off threshold outlined in the published manual (Tombaugh, 1996).

Data collection

Demographic, injury, financial compensation, pre-injury and post-injury medical status predictors of interest were extracted during file review from neuropsychological reports, the Liverpool Hospital BIRU patient database, and Liverpool Hospital medical records.

Predictor variables

Demographics

Sex, age at assessment, years of education, and working at time of injury (yes/no) were included as predictor variables.

Culturally and linguistically diverse (CALD) status

Relevant data collected included country/region of birth, language(s) spoken, country/language in which the case was educated, and whether an interpreter was used during the neuropsychological

assessment. CALD status was defined as a four-level variable, as follows: (1) English-speaking background (ESB) and educated at an English-speaking educational institution; (2) CALD background, as defined by: a first language other than English (including bilingual cases who later became proficient in English) or born overseas in a country where English is not the official language, and educated predominantly at an English-speaking educational institution (CALD-E); (3) CALD and educated at an institution where English was not the predominant language, but did not require an interpreter during the neuropsychological assessment (CALD-N); and (4) CALD and required an interpreter during the assessment (CALD-I) (Coward, 2012).

Injury-related variables

Time since injury (months), cause of injury, and injury severity data (initial Glasgow coma scale (GCS) and post-traumatic amnesia (PTA) duration) were collected. For 27 cases, PTA duration was missing but a reasonable estimate could be made based on related clinical data. For an additional 34 cases where this information was not available, PTA duration was left as missing data. TBI severity was defined according to the Department of Veterans Affairs/Department of Defense consensus-based classification of closed TBI severity – mild: <1 day PTA, moderate: 1–7 days PTA, and severe: >7 days PTA (Department of Veterans Affairs, 2009).

Financial compensation

Financial compensation status was entered as a dichotomous (yes/no) predictor variable. Cases classified as having a positive financial compensation status were either: a participant at the time

Table 1. Assessment and demographic characteristics of the final sample and according to TOMM group

	<i>n</i>	Final Sample (<i>N</i> = 491)	TOMM Pass (<i>n</i> = 443)	TOMM Fail (<i>n</i> = 48)	$\chi^2/z/t$	<i>p</i>	<i>p</i> (FDR)	Effect Size OR/ <i>r</i> / <i>d</i>	95%CI	
Setting	491				21.93	<.0001	<.0004	6.42	2.68	15.42
Inpatient		218 (44.40%)	212 (97.25%)	6 (2.75%) ^a						
Outpatient		273 (55.60%)	231 (84.62%)	42 (15.38%) ^b						
Age at Assessment Med (IQR) Range	491	40 (26–52) 16–73	39 (25–52)	41.5 (29–54)	1.02 ^c	.31	.38	0.05		
Education (years) M (SD) Range	491	11.29 (2.77) 0–20	11.30 (2.71)	11.15 (3.32)	–0.31 ^d	.76	.81	0.05	–0.29	0.40
Sex	491				0.41	.52	.59	1.25	0.63	2.50
Male		386 (78.62%)	350 (90.67%)	36 (9.33%)						
Female		105 (21.38%)	93 (88.57%)	12 (11.43%)						
Working at Injury Yes No ^e	477	332 (69.60%) 145 (30.40%)	295 (88.86%) 136 (93.79%)	37 (11.14%) 9 (6.21%)	2.82	.09	.18	1.90	0.89	4.04

Note. Effect sizes are odds ratios for χ^2 tests, *r* for Wilcoxon rank-sum tests, and Cohen's *d* for independent samples *t*-tests. CI = confidence interval; FDR = false discovery rate; OR = odds ratio; TOMM = test of memory malingering.

^aFor inpatient cases that failed the TOMM, scores ranged from 37–44/50 on Trial 2 and 41–50/50 on the Retention Trial.

^bFor outpatient cases that failed the TOMM, scores ranged from 14–44/50 on Trial 2 and 3–50/50 on the Retention Trial.

^cZ-score corresponding to Wilcoxon rank-sum *W_s*.

^dIndependent samples *t*-test.

^eIncludes cases that were studying or had retired, in addition to unemployed cases.

of assessment in one of the three main “no-fault” insurance programs available in New South Wales for persons who sustained a TBI in a motor vehicle or workplace accident; potentially eligible and currently had an application under review for acceptance into an insurance program at assessment; or were subsequently accepted into an insurance program within one year of assessment.

Mental health variables

Three mental health variables were defined dichotomously (yes/no): pre-injury drug and alcohol history, pre-injury psychological history, and post-injury psychological distress.

A positive pre-injury drug and alcohol history indicated any lifetime diagnosis of alcohol or substance use disorder, patient or family reports of heavy and prolonged drinking/substance use that interfered with their usual activities of daily living, or treatment for alcohol or substance use issues (e.g., regular counseling, substitution/withdrawal programs, inpatient psychiatric/rehabilitation admission).

A positive pre-injury psychological history reflected any lifetime diagnosis of affective or other mental health disorder, or prior intervention with a mental health professional or medication for treatment of psychological symptoms before injury.

Post-injury psychological distress was recorded as positive if the case reported experiencing significant psychological distress at assessment based on clinical interview and formal psychological inventories (e.g., the Depression, Anxiety and Stress Scales (DASS)), or if they continued to meet formal criteria for a mental health disorder or receive a mental health intervention (including psychotropic medication) at some point after their injury that had not resolved by the time of assessment. DASS subscale scores were not examined separately as the DASS was not administered uniformly throughout the sample (71% completion rate).

Statistical analysis

IBM SPSS Statistics v26.0 and JMP v14.2 were used to conduct the statistical analysis. Descriptive statistics were generated for all variables and univariate comparisons of the TOMM pass/fail groups

were conducted using χ^2 , independent samples *t*-tests, or Wilcoxon rank-sum tests as appropriate. A false discovery rate (FDR) correction for multiple comparisons (Benjamini & Hochberg, 1995) was applied to reported *p*-values.

Next, statistically significant predictors at the univariate level (*p*(FDR) < .05, two-sided) were entered into a forward stepwise nominal logistic regression model with TOMM classification as the outcome variable. Additional follow-up models (each containing two main effect terms corresponding to a pair of significant multivariate predictors identified from the previous model as well as their interaction) were generated to further explore whether any mediating or moderating variables could be identified. Effect sizes (odds ratios, Cohen's *d*, and *r*) and 95% confidence intervals were calculated for all relevant comparisons.

Results

Sample characteristics

A summary of the final sample (*n* = 491) descriptive statistics can be found across Tables 1–4. Of note, inpatient (44%) and outpatient assessments (56%) were both well represented. The sample was predominantly male (79%) with an average of 11 years of formal education. One-third were classified as CALD (of which 44% were CALD-I). The vast majority (82%) sustained a severe TBI (>7 days PTA). Just under half (49%) had a positive financial compensation status. One quarter (28%) had a pre-injury psychological history and the same proportion continued to experience psychological distress post-injury.

The excluded cases (*n* = 85) were not significantly different from the final sample in relation to assessment setting, age, sex, education, CALD status, injury cause, or financial compensation status. In terms of injury severity, while the excluded cases had longer median PTA duration than the final sample (Med = 34 days vs. 24 days; *p*(FDR) = .002), both groups overwhelmingly represented severely injured cases (89% vs. 82%; *p*(FDR) = .43). Excluded cases were also more likely to report psychological distress post-injury (42% vs. 28%; *p*(FDR) = .03).

Table 2. CALD status of the final sample and according to TOMM group

	<i>n</i>	Final Sample (<i>N</i> = 491)	TOMM Pass (<i>n</i> = 443)	TOMM Fail (<i>n</i> = 48)	χ^2	<i>p</i>	<i>p</i> (FDR)	OR	95%CI
CALD Status	491				51.91 ^a	<.0001	<.0004		
ESB		326 (66.40%)	312 (95.71%)	14 (4.29%)					
CALD-E		62 (12.63%)	54 (87.10%)	8 (12.90%)					
CALD-N		30 (6.11%)	27 (90.30%)	3 (9.70%)					
CALD-I		73 (14.87%)	50 (68.49%)	23 (31.51%)					
Interpreter Required?	491				45.91	<.0001	<.0004	7.23	3.82 13.69
Yes		73 (14.87%)	50 (68.49%)	23 (31.51%)					
No		418 (85.13%)	393 (94.02%)	25 (5.98%)					

Note. CALD-E = culturally and linguistically diverse-English educated; CALD-N = culturally and linguistically diverse-non-English educated; CALD-I = culturally and linguistically diverse-interpreter; CI = confidence interval; ESB = English-speaking background; FDR = false discovery rate; OR = odds ratio; TOMM = test of memory malingering.
^aAnalysis of standardized residuals confirmed that the association between TOMM failure and CALD status was predominantly driven by the CALD-I sub-group ($z = 6.02, p(\text{FDR}) < .0008$) with fewer than expected ESB cases failing the TOMM ($z = -3.17, p(\text{FDR}) = .0008$). The number of CALD-E and CALD-N cases that failed the TOMM remained within expectations ($|z| < .79, p(\text{FDR}) > .40$).

Table 3. Injury characteristics of the final sample and according to TOMM group

	<i>n</i>	Final Sample (<i>N</i> = 491)	TOMM Pass (<i>n</i> = 443)	TOMM Fail (<i>n</i> = 48)	χ^2	<i>p</i>	<i>p</i> (FDR)	Effect Size OR/ <i>r</i>	95%CI
Injury Cause	481				0.57	.90	.90		
On-Road Accident ^a		249 (51.77%)	225 (90.36%)	24 (9.64%)					
Fall		142 (29.52%)	128 (90.14%)	14 (9.86%)					
Assault		55 (11.43%)	48 (87.27%)	7 (12.73%)					
Other ^b		35 (7.28%)	32 (91.43%)	3 (8.57%)					
GCS Med (IQR)	471	10 (4–14)	9 (4–14)	13 (5.5–14)	1.54 ^c	.12	.21	0.07	
GCS Categories	471				5.75	.06	.14	0.44 ^d	0.22 0.87
3 to 8		223 (47.34%)	209 (93.72%)	14 (6.28%)					
9 to 12		59 (12.53%)	53 (89.83%)	6 (10.17%)					
13 to 15		189 (40.13%)	164 (86.77%)	25 (13.23%)					
PTA Med (IQR)	457	24 (10–39)	24 (10–39)	24 (7–40.5)	-1.09 ^c	.27	.36	-0.05	
TBI Severity	463				10.34	.006	.02	4.65 ^e	1.68 12.87
Mild (< 1 day PTA)		21 (4.54%)	15 (71.43%)	6 (28.57%)					
Moderate (1–7 days PTA)		65 (14.04%)	59 (90.77%)	6 (9.23%)					
Severe (>7 days PTA)		379 (81.86%)	349 (92.08%)	30 (7.92%)					
Financial Compensation	491				3.95	.05	.13	1.85	1.01 3.45
Compensable		240 (48.88%)	210 (87.50%)	30 (12.50%)					
Noncompensable		251 (51.12%)	233 (92.83%)	18 (7.17%)					

Note. Effect sizes are odds ratios for χ^2 tests and *r* for Wilcoxon rank-sum tests. CI = confidence interval; FDR = false discovery rate; GCS = Glasgow coma scale; OR = odds ratio; PTA = post-traumatic amnesia; TOMM = test of memory malingering.
^aOn-road accident includes motor vehicle accidents (MVAs) (*n* = 123), motorbike accidents (*n* = 56), pedestrian vs. MVAs (*n* = 50), and pushbike accidents (*n* = 20).
^bOther includes blast/penetrating injury (*n* = 11), crush injury (*n* = 3), sports-related injury (*n* = 10), blunt force from falling/moving object (*n* = 6), skateboarding accidents (*n* = 4), pedestrian vs. train accidents (*n* = 2), unspecified work-related injury (*n* = 1), plane crash (*n* = 1), boating accident (*n* = 1).
^cZ-score corresponding to Wilcoxon rank-sum W_s .
^dOR corresponding to 2x2 contingency table comparing GCS = 3–8 to GCS = 13–15.
^eOR corresponding to 2x2 contingency table comparing <1 day PTA to >7 days PTA.

Comparison of TOMM pass/fail groups

In total, 48 (9.8%) of 491 cases failed the TOMM. Of those that failed, the median scores were 29/50 (IQR = 23–37) on Trial 1, 37/50 (IQR = 27–42) on Trial 2, and 37/50 (IQR = 27.75–44) on the Retention Trial. Five cases (1%) scored well below chance (<18/50) on Trial 1, as did one (0.2%) on Trial 2 and two (0.4%) on the Retention Trial.

Univariate comparisons of the predictor variables according to TOMM outcome are displayed across Tables 1–4. Significant predictors of TOMM failure were: outpatient setting, CALD-I, mild TBI (<1 day PTA), and post-injury psychological distress ($p(\text{FDR}) < .02$). Other demographic, injury cause, GCS classification, financial compensation status, and mental health variables were unrelated to TOMM performance.

Of note, there was an inverse dose-response relationship between injury severity and TOMM failure, with 6 of 21 (28%) mild TBI cases

failing the TOMM, compared with only 6 of 65 (9%) moderate and 30 of 379 (8%) severe cases. At the other end of the injury spectrum, seven of nine (78%) cases with PTA duration >6 months (reflecting the most disabling injuries given PTA at this duration is associated with chronic amnesic syndrome) successfully passed the TOMM. The two failures were just below the conventional cut-off on Trial 2 (both scoring 44/50) and were able to pass the Retention Trial (scoring 45/50 and 50/50 respectively).

Logistic regression model

Results of the logistic regression analysis of TOMM failure are presented in Table 5. Predictors considered for forward stepwise selection were: assessment setting, TBI severity, CALD-I, and post-injury psychological distress. There was no evidence of significant collinearity among the predictors (Cramer’s $V_s < .35$). Overall, the final model -2 log likelihood ratio was significant ($\chi^2 = 58.81$,

Table 4. Mental health histories of the final sample and according to TOMM group

	<i>n</i>	Final Sample (<i>n</i> = 491)	TOMM Pass (<i>n</i> = 43)	TOMM Fail (<i>n</i> = 48)	χ^2	<i>p</i>	<i>p</i> (FDR)	OR	95%CI
Pre-Injury Drug & Alcohol History	491				2.12	.15	.22	0.52	0.22 1.27
Yes		101 (20.57%)	95 (94.06%)	6 (5.94%)					
No		390 (79.43%)	348 (89.23%)	42 (10.77%)					
Pre-Injury Psychological History	491				2.13	.14	.22	0.57	0.27 1.22
Yes		136 (27.70%)	127 (93.38%)	9 (6.62%)					
No		355 (72.30%)	316 (89.01%)	39 (10.99%)					
Post-Injury Psychological Distress	491				15.14	<.0001	<.0004	3.17	1.73 5.81
Yes		138 (28.11%)	113 (81.88%)	25 (18.12%)					
No		353 (71.89%)	330 (93.48%)	23 (6.52%)					

Note. CI = confidence interval; FDR = false discovery rate; OR = odds ratio; TOMM = test of memory malingering.

Table 5. Logistic regression analysis

	Nagelkerke $R^2/\Delta R^2$	β (SE)	Wald	<i>p</i>	<i>p</i> (FDR)	aOR	95%CI
Step 1	.14****						
Constant		-2.84 (0.20)	167.46	<.0001	<.0001	0.06	
Interpreter [No = 0]		1.98 (0.35)	32.97	<.0001	<.0001	7.27	3.70 14.32
Step 2	.09****						
Constant		-4.14 (0.46)	81.91	<.0001	<.0001	0.02	
Interpreter [No = 0]		2.04 (0.36)	31.66	<.0001	<.0001	7.72	3.79 15.74
Setting [Inpatient = 0]		1.83 (0.47)	15.41	<.0001	<.0001	6.24	2.50 15.55
Step 3	.03**						
Constant		-4.36 (0.48)	82.84	<.0001	<.0002	0.01	
Interpreter [No = 0]		2.11 (0.38)	31.73	<.0001	<.0002	8.25	3.96 17.18
Setting [Inpatient = 0]		1.57 (0.48)	10.66	.001	.0013	4.80	1.87 12.31
Post-Injury Psychological Distress [No = 0]		1.02 (0.37)	7.70	.006	.006	2.77	1.35 5.70

Note. aOR = adjusted odds ratio; CI = confidence interval; FDR = false discovery rate.

p* < .01; **p* < .0001.

p < .0001) with Nagelkerke $R^2 = .26$, indicating that the model explained around 26% of variance in TOMM outcome.

Three of the four variables were included in the final multivariate model. By far the strongest predictor was CALD-I (*p*(FDR) < .0002), which was associated with an adjusted odds ratio (aOR) of 8.25-fold greater likelihood of TOMM failure relative to non-interpreter cases. The other significant predictors in the model were outpatient setting (*p*(FDR) = .001; aOR = 4.80 vs. inpatient setting) and post-injury psychological distress (*p*(FDR) = .006; aOR = 2.77).

Further examination of multivariate predictors of TOMM failure

Three additional models were conducted to further explore associations between the three multivariate predictors, each of which contained two main effects corresponding to two of the three significant multivariate predictors and their interaction (i.e., Model 1 = CALD-I + assessment setting + CALD-I \times assessment setting; Model 2 = CALD-I + post-injury psychological distress + CALD-I \times post-injury psychological distress; Model 3 = assessment setting + post-injury psychological distress + assessment setting \times post-injury psychological distress). Each of these models yielded nonsignificant interaction effects (*p*(FDR)s > .82). This outcome was not entirely unexpected as the models were considerably underpowered on account of very low expected and observed frequencies (<5) in the TOMM failure \times inpatient cells.

Nonetheless, there was a strong indication that TOMM performance varied for CALD-I cases depending on the assessment

setting (see Figure 2). Specifically, 21 of 43 (49%) outpatient CALD-I cases failed the TOMM compared with only 2 of 30 (7%) inpatient CALD-I cases ($\chi^2 = 14.56$, *p*(FDR) = .0006; OR = 13.36, 95%CI: 2.82–63.23). By comparison, 21 of 230 (9%) non-interpreter outpatients and 4 of 188 (2%) non-interpreter inpatients failed the TOMM ($\chi^2 = 9.02$, *p*(FDR) = .006; OR = 4.62, 95%CI: 1.56–13.71).

The relationship between post-injury psychological distress and TOMM failure did not appear to be mediated by interpreter use. CALD-I cases were more likely to fail the TOMM when they concurrently presented with psychological distress (11 of 23; 48%) than when they did not (12 of 50; 24%) ($\chi^2 = 4.14$, *p*(FDR) = .05; OR = 2.90, 95%CI: 1.02–8.25). However, a similar effect size was observed in the non-interpreter cases, with 14 of 115 (12%) presenting with current psychological distress failing the TOMM compared with 11 of 303 (4%) that did not present as distressed ($\chi^2 = 10.82$, *p*(FDR) = .003; OR = 3.68, 95%CI: 1.62–8.37).

It is important to emphasize that the CALD-I cases were no more likely to be assessed as outpatients (59% vs. 55%) or to present with post-injury psychological distress (32% vs. 28%) than non-interpreter cases. Additionally, there was no difference in education levels between the CALD-I (*M* = 10.9 years) and non-interpreter cases (*M* = 11.4 years).

The interaction between assessment setting and post-injury psychological distress in predicting TOMM failure could not be analyzed as an odds ratio for the psychological distress/TOMM failure relationship could not be calculated at the inpatient level. This was because none of the 25 inpatients presenting with

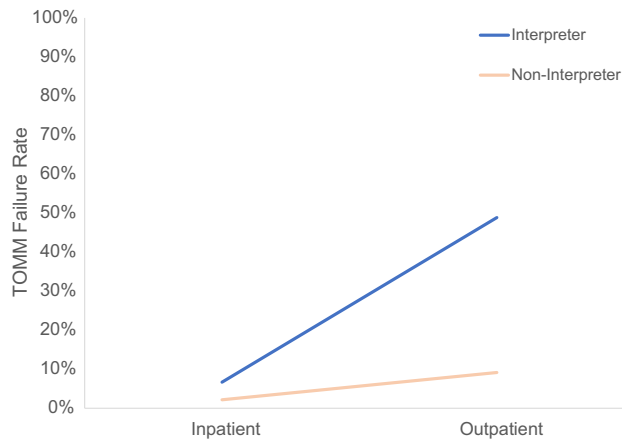


Figure 2. TOMM failure rate – interaction between CALD-I and assessment setting.

psychological distress failed the TOMM. Six of 193 (3%) inpatient cases that did not report psychological distress failed ($\chi^2 = 0.80$, $p(\text{FDR}) = .37$). In the outpatient setting, 25 of 113 (22%) cases presenting with psychological distress failed the TOMM, compared with 17 of 160 (11%) that did not ($\chi^2 = 6.73$, $p(\text{FDR}) = .02$; $\text{OR} = 2.39$, 95%CI: 1.22–4.67).

Discussion

To our knowledge, this study reflects the largest examination to date of TOMM performance in a rehabilitation setting predominantly focused on managing TBI patients at the more severe end of the injury spectrum (82% had a severe injury as defined by PTA duration >7 days). We found that 9.8% of a consecutive sample ($N = 491$) failed the TOMM. The study identified several novel predictors of TOMM failure, most notably interpreter use and outpatient setting. Post-injury psychological distress was also associated with TOMM failure. Importantly, nearly half of CALD-I cases failed the TOMM when assessed in the outpatient setting, compared with only 7% of inpatient CALD-I cases.

The nearly 10% rate of TOMM failure in the current study was very similar to that reported in a smaller study of consecutively referred TBI rehabilitation patients (Moore & Donders, 2004). This was despite there being a much higher proportion of severely injured cases, a lower proportion of outpatient assessments, and much greater cultural diversity in the current sample, which attests to the robustness of this finding within the brain injury rehabilitation setting. It is also in line with the base rates reported in many smaller studies of other clinical groups, including stroke (Bodner et al., 2019), Korsakoff's syndrome (Oudman et al., 2020), Huntington's disease (Sieck et al., 2012), mild cognitive impairment (Walter et al., 2014), and HIV (Paul et al., 2017).

An inverse dose-response relationship between injury severity and PVT failure was observed in the current study, with TOMM failure rates being much higher in mild TBI cases compared to moderate and severe cases. This paradoxical finding of higher PVT failure rates in mild TBI cases is well established in the PVT literature in both medico-legal (e.g., Constantinou et al., 2005; Green et al., 2001; Larrabee, 2007) and clinical TBI studies (e.g., Sherry et al., 2021), and has generally been found to be mediated by secondary gain incentives. That said, there were only a small number of mild TBI cases in this study (<5% of the study sample) and this sub-group was over-represented by both

outpatient and CALD-I cases. We therefore caution that they are likely to be unrepresentative of the broader clinical mild TBI population.

Our analysis emphasizes that assessment setting has an important influence on TOMM outcome. Overall, 15% of outpatients failed the TOMM, similar to the 21.8% failure rate reported in another outpatient TBI rehabilitation sample (Locke et al., 2008). The very low rate of TOMM failure seen in inpatients (2.8%) is also consistent with clinical studies that used other PVTs in other inpatient populations (e.g., Gierok et al., 2005; Lee et al., 2021).

We propose several potential reasons for the disparity in TOMM failure rates according to the assessment setting. On the one hand, inpatients may have a greater incentive to perform well on neuropsychological testing if the assessment is presented to them as an opportunity to demonstrate the extent of their recovery, and they believe that a strong performance may bolster their case for an earlier discharge. On the other hand, after returning to the community, patients may be increasingly exposed to factors relating to secondary gain (e.g., legal claims, disability payments, access to additional supports and services, or pressure to return to unwanted pre-injury responsibilities such as work or study). There could also be a heightened focus at this time on primary gain factors such as psychological maladjustment and related symptoms (e.g., pain, fatigue, headaches, sleep disturbance), leading patients to perform poorly on PVTs as a way of expressing their distress to their healthcare providers. Such possibilities require further investigation.

In the current study, CALD background did not increase the risk of TOMM failure in and of itself. Rather, the driving factor was the need for an interpreter, particularly in the outpatient setting. The differential rates of TOMM failure in CALD-I cases across inpatient (7%) and outpatient (49%) settings indicate that factors such as the presence of the interpreter in the room or the accuracy of translation of test instructions did not significantly influence TOMM performance. Interpreter use may be a proxy for any number of socio-cultural and language factors, including limited English proficiency, education style and quality, acculturation effects, and immigration context (Fujii, 2018), that we were unable to analyze in greater detail in this study. For example, higher failure rates in CALD-I cases may reflect differences in coping styles, beliefs regarding their level of personal control over their recovery, and broader social or cultural norms regarding appropriate expression of the sick role, particularly after returning to the community (Maffini & Wong, 2014; Ponsford et al., 2020). Moreover, CALD-I cases may not have equitable access to relevant health or rehabilitation information, with limited access to interpreters in community settings (Harrison et al., 2020). Alternatively, the mismatch of cultural backgrounds between the patient and the examiner could be impacting (Barber Rioja & Rosenfeld, 2018), although such an explanation is undermined by the differential effect of assessment setting on TOMM failure in CALD-I cases in the current study. Such considerations reflect the broader challenges in improving health literacy within CALD communities and require further investigation with regard to their relevance to TOMM outcome.

Additionally, we found that education did not influence TOMM performance in the CALD-I sub-group. A handful of previous studies (e.g., Nijdam-Jones et al., 2017) have reported low TOMM scores in individuals from non-English-speaking countries and suggested that education may be a mediating factor. However, we note that the CALD-I sub-group completed an average of 11 years of school, in line with the rest of the sample. It may

be the case that a strong inverse relationship between education and TOMM failure is apparent only in those with limited formal education (e.g., <6 years) and/or literacy skills (Nijdam-Jones et al., 2019).

Our analysis of the CALD sub-group has broader implications for PVT research. While some studies (e.g., Salazar et al., 2007) have argued that alternate TOMM cut-off scores are needed for different cultural groups, applying the conventional cut-off scores resulted in only 7% of CALD inpatients in the current study failing the TOMM, suggesting that they are appropriate for use with CALD individuals following TBI (also see Martin et al., 2020). On the other hand, the very high rate of TOMM failure in CALD-I outpatients raises concerns over the validity of past research that has compared neuropsychological performance in CALD individuals with their English-speaking counterparts, as it suggests that if PVTs were not administered, poorer performance in CALD groups may be over-estimated due to the inadvertent inclusion of some non-credible performers.

The presence of psychological distress at the time of assessment was also uniquely related to TOMM failure in the current study. Psychological functioning has been inconsistently associated with PVT failure in the clinical literature (e.g., Ashendorf et al., 2004; Locke et al., 2008; Moore & Donders, 2004; Rai & Erdodi, 2021; Rees et al., 2001) which is likely to reflect differences in the medical condition(s) being examined as well as how psychological distress has been defined across studies. In the current study, we were unable to determine the validity of the reported psychological distress in these cases as symptom validity tests were not included as part of routine clinical assessments. Inclusion of such measures will be an important consideration for future research to untangle the unique contribution of genuine psychological distress from the possibility of exaggeration or symptom feigning (Donders, Lefebvre, & Goldsworthy, 2021).

It is interesting to note that we did not find an association between financial compensation-seeking and TOMM failure after correcting for multiple comparisons. This result stands in contrast with much of the PVT literature, although a predictive relationship has been reported more commonly in medico-legal, forensic, and veteran samples compared to clinical samples (Larrabee, 2007). The disparity may also partly relate to differences in the nature of the financial compensation being sought, which for most cases consisted of a “no-fault” lifetime insurance program to cover ongoing medical, rehabilitation, and other care expenses, as opposed to a lump sum payment that may be present in other jurisdictions, and that is more typical of the underlying secondary gain incentive present in medico-legal PVT studies.

The current study has some important clinical implications that are particularly pertinent to the Australian context in light of Uiterwijk et al.'s (2021) recent report of inconsistent PVT use among Australian clinicians. First, the study findings strongly advocate for routinely administering PVTs in clinical assessments of TBI patients, particularly in outpatient settings. We recognize that failure on a single PVT can occur in credible patients (Victor et al., 2009) and therefore the TOMM outcome should not be wholly relied upon to assess performance validity. Rather, it needs to be considered alongside multiple other validated PVTs to optimize overall diagnostic accuracy (Sherman et al., 2020; Sweet, Heilbronner, et al., 2021). Second, the current findings clearly demonstrate that the TOMM can be validly administered to most inpatients with severe TBI. Moreover, our experience is that once these patients have emerged from PTA, they are capable

of undertaking a comprehensive neuropsychological assessment (usually spread over multiple shorter sessions). Based on our results, we argue that there are advantages to conducting initial neuropsychological assessments of TBI patients while they are inpatients, insofar as the chances of obtaining valid test results are likely to be higher.

This study has several limitations, first and foremost being its retrospective nature. Beyond deciding to systematically administer the TOMM, we did not seek to interfere with any other aspects of routine clinical practice. While scores on other embedded PVTs were recorded, they were administered inconsistently according to clinical need, and we were therefore unable to use them to generate a second criterion for classifying invalid responding. Similarly, while neuropsychological test data were collected, we were also unable to collate a common set of measures that could be used to compare cognitive performance among the different sub-groups examined in this study. Our reliance on archival records also impacted the scope and quality of available information for some variables (particularly relating to CALD status and mental health history), preventing a more sophisticated analysis. It will be important to consider each of these issues prospectively in future studies to be able to explore whether alternate TOMM cut-offs provide better sensitivity and specificity in more severely injured TBI patients, to better understand the factors underpinning the social and cultural influences on TOMM performance observed in the current sample, and to compare TOMM performance against other “culturally neutral” PVTs in CALD-I individuals. Second, the retrospective study design also prevented us from discerning the likely motivations for TOMM failure within this sample. Third, the excluded cases reported higher rates of post-injury psychological distress. This raises the possibility of a selection bias leading us to slightly underestimate the true base rate of TOMM failure in this severely injured population. We would however apply a large caveat to this statement given the very general operationalization of the psychological distress variable in this study along with the aforementioned concerns over symptom validity. Finally, as noted above, our analysis of the interactions between multivariate predictors and TOMM outcome were considerably underpowered on account of the small expected and observed frequencies in the inpatient/TOMM failure cells. To optimally determine whether any of these interactions reflect real and clinically meaningful effects, future research should consider settings where large samples with a higher ratio of inpatient to outpatient cases can be studied.

In conclusion, our findings strongly reinforce several key points outlined in the recent update to the 2009 AACN Consensus Statement (Sweet, Heilbronner, et al., 2021). The results add to a growing literature indicating that PVT failure does occur in routine clinical settings and is associated with multiple factors beyond the potential for secondary gain. They strongly support the case for administering PVTs in clinical assessment of TBI patients, rather than relying on clinical judgment, to ensure that neuropsychological test results can be interpreted with a high degree of confidence.

Acknowledgements. The authors would like to thank Dr Adeline Hodgkinson (Director, Liverpool Hospital BIRU) for her support on this project as well as other BIRU neuropsychologists who conducted assessments that were included in the study. They would also like to thank Dr Wei Xuan (Senior Biostatistician, Ingham Institute for Applied Medical Research) for statistical advice regarding the logistic regression modeling used in the study.

Funding statement. This work was supported by a South-Western Sydney Local Health District Allied Health Seeding Grant (2018) awarded to Thomas Gates. Partial salary support for Kate Young was provided through the Ingham Institute for Applied Medical Research (Brain Injury Rehabilitation Research Group).

Conflicts of interest. None.

References

- Ashendorf, L., Constantinou, M., & McCaffrey, R. J. (2004). The effect of depression and anxiety on the TOMM in community-dwelling older adults. *Archives of Clinical Neuropsychology*, *19*(1), 125–130. <https://doi.org/10.1093/arclin/19.1.125>
- Australian Bureau of Statistics. (2017). 2016 Census QuickStats – Sydney – South West. Retrieved from https://quickstats.censusdata.abs.gov.au/census_services/getproduct/census/2016/quickstat/127?opendocument
- Barber Rioja, V., & Rosenfeld, B. (2018). Addressing linguistic and cultural differences in the forensic interview. *The International Journal of Forensic Mental Health*, *17*(4), 377–386. doi: [10.1080/14999013.2018.1495280](https://doi.org/10.1080/14999013.2018.1495280)
- Bashem, J. R., Rapport, L. J., Miller, J. B., Hanks, R. A., Axelrod, B. N., & Millis, S. R. (2014). Comparisons of five performance validity indices in bona fide and simulated traumatic brain injury. *The Clinical Neuropsychologist*, *28*(5), 851–875. doi: [10.1080/13854046.2014.927927](https://doi.org/10.1080/13854046.2014.927927)
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *57*(1), 289–300.
- Bodner, T., Merten, T., & Benke, T. (2019). Performance validity measures in clinical patients with aphasia. *Journal of Clinical and Experimental Neuropsychology*, *41*(5), 476–483. doi: [10.1080/13803395.2019.1579783](https://doi.org/10.1080/13803395.2019.1579783)
- Constantinou, M., Bauer, L., Ashendorf, L., Fisher, J. M., & McCaffrey, R. J. (2005). Is poor performance on recognition memory effort measures indicative of generalized poor performance on neuropsychological tests? *Archives of Clinical Neuropsychology*, *20*(2), 191–198. doi: [10.1016/j.acn.2004.06.002](https://doi.org/10.1016/j.acn.2004.06.002)
- Coward, B. (2012). *Predictors of cognitive outcome following traumatic brain injury: A retrospective study.* (Unpublished doctoral dissertation). Macquarie University, Sydney, Australia.
- Department of Veteran's Affairs. (2009). VA/DoD clinical practice guideline for management of concussion/mild traumatic brain injury. *Journal of Rehabilitation Research and Development*, *46*(6), CP1–CP68. doi: [10.1682/JRRD.2009.06.0076](https://doi.org/10.1682/JRRD.2009.06.0076)
- Donders, J., Lefebvre, N., & Goldsworthy, R. (2021). Patterns of performance and symptom validity test findings after mild traumatic brain injury. *Archives of Clinical Neuropsychology*, *36*(3), 394–402. doi: [10.1093/arclin/acz057](https://doi.org/10.1093/arclin/acz057)
- Fox, D. D. (2011). Symptom validity test failure indicates invalidity of neuropsychological tests. *The Clinical Neuropsychologist*, *25*(3), 488–495. doi: [10.1080/13854046.2011.554443](https://doi.org/10.1080/13854046.2011.554443)
- Fujii, D. E. M. (2018). Developing a cultural context for conducting a neuropsychological evaluation with a culturally diverse client: The ECLECTIC framework. *The Clinical Neuropsychologist*, *32*(8), 1356–1392. doi: [10.1080/13854046.2018.1435826](https://doi.org/10.1080/13854046.2018.1435826)
- Gierok, S. D., Dickson, A. L., & Cole, J. A. (2005). Performance of forensic and non-forensic adult psychiatric inpatients on the Test of Memory Malingering. *Archives of Clinical Neuropsychology*, *20*(6), 755–760. doi: [10.1016/j.acn.2005.04.008](https://doi.org/10.1016/j.acn.2005.04.008)
- Green, P., Rohling, M. L., Lees-Haley, P. R., & Iii, L. M. A. (2001). Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury*, *15*(12), 1045–1060. doi: [10.1080/02699050110088254](https://doi.org/10.1080/02699050110088254)
- Greve, K. W., Bianchini, K. J., & Doane, B. M. (2006). Classification accuracy of the test of memory malingering in traumatic brain injury: Results of a known-groups analysis. *Journal of Clinical and Experimental Neuropsychology*, *28*(7), 1176–1190. doi: [10.1080/13803390500263550](https://doi.org/10.1080/13803390500263550)
- Harrison, R., Walton, M., Chitkara, U., Manias, E., Chauhan, A., Latanik, M., & Leone, D. (2020). Beyond translation: Engaging with culturally and linguistically diverse consumers. *Health Expectations*, *23*(1), 159–168. doi: [10.1111/hex.12984](https://doi.org/10.1111/hex.12984)
- Johnson-Greene, D., Brooks, L., & Ference, T. (2013). Relationship between performance validity testing, disability status, and somatic complaints in patients with fibromyalgia. *The Clinical Neuropsychologist*, *27*(1), 148–158. doi: [10.1080/13854046.2012.733732](https://doi.org/10.1080/13854046.2012.733732)
- Kemp, S., Coughlan, A. K., Rowbottom, C., Wilkinson, K., Teggart, V., & Baker, G. (2008). The base rate of effort test failure in patients with medically unexplained symptoms. *Journal of Psychosomatic Research*, *65*(4), 319–325. doi: [10.1016/j.jpsychores.2008.02.010](https://doi.org/10.1016/j.jpsychores.2008.02.010)
- Larrabee, G. J. (2007). Introduction: Malingering, research designs, and base rates. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 3–13). New York, NY: Oxford University Press.
- Larrabee, G. J., Millis, S. R., & Meyers, J. E. (2009). 40 plus or minus 10, a new magical number: Reply to Russell. *The Clinical Neuropsychologist*, *23*(5), 841–849. doi: [10.1080/13854040902796735](https://doi.org/10.1080/13854040902796735)
- Lee, B. G., Kent, J.-A., Marcopulos, B. A., Arredondo, B. C., & Wilson, M. (2021). Rey–Osterrieth complex figure normative data for the psychiatric population. *The Clinical Neuropsychologist*, Advance online publication. doi: [10.1080/13854046.2021.1897163](https://doi.org/10.1080/13854046.2021.1897163)
- Locke, D. E. C., Smigielski, J. S., Powell, M. R., & Stevens, S. R. (2008). Effort issues in post-acute outpatient acquired brain injury rehabilitation seekers. *NeuroRehabilitation*, *23*(3), 273–281.
- Maffini, C. S., & Wong, Y. J. (2014). Assessing somatization with Asian American clients. In L. T. Benuto, L. T. Thaler, & B. D. Leany (Eds.), *Guide to psychological assessment with Asians* (pp. 347–360). New York, NY: Springer.
- Martin, P. K., & Schroeder, R. W. (2020). Base rates of invalid test performance across clinical non-forensic contexts and settings. *Archives of Clinical Neuropsychology*, *35*(6), 717–725. doi: [10.1093/arclin/acia017](https://doi.org/10.1093/arclin/acia017)
- Martin, P. K., Schroeder, R. W., & Odland, A. P. (2015). Neuropsychologists' validity testing beliefs and practices: A survey of North American professionals. *The Clinical Neuropsychologist*, *29*(6), 741–776. doi: [10.1080/13854046.2015.1087597](https://doi.org/10.1080/13854046.2015.1087597)
- Martin, P. K., Schroeder, R. W., Olsen, D. H., Maloy, H., Boettcher, A., Ernst, N., & Okut, H. (2020). A systematic review and meta-analysis of the Test of Memory Malingering in adults: Two decades of deception detection. *The Clinical Neuropsychologist*, *34*(1), 88–119. doi: [10.1080/13854046.2019.1637027](https://doi.org/10.1080/13854046.2019.1637027)
- McWhirter, L., Ritchie, C. W., Stone, J., & Carson, A. (2020). Performance validity test failure in clinical populations—a systematic review. *Journal of Neurology, Neurosurgery, & Psychiatry*, *91*(9), 945–952. doi: [10.1136/jnnp-2020-323776](https://doi.org/10.1136/jnnp-2020-323776)
- Moore, B. A., & Donders, J. (2004). Predictors of invalid neuropsychological test performance after traumatic brain injury. *Brain Injury*, *18*(10), 975–984. doi: [10.1080/02699050410001672350](https://doi.org/10.1080/02699050410001672350)
- Nijdam-Jones, A., Rivera, D., Rosenfeld, B., & Arango-Lasprilla, J. C. (2017). A cross-cultural analysis of the Test of Memory Malingering among Latin American Spanish-speaking adults. *Law and Human Behavior*, *41*(5), 422–428. doi: [10.1037/lhb0000250](https://doi.org/10.1037/lhb0000250)
- Nijdam-Jones, A., Rivera, D., Rosenfeld, B., & Arango-Lasprilla, J. C. (2019). The effect of literacy and culture on cognitive effort test performance: An examination of the Test of Memory Malingering in Colombia. *Journal of Clinical and Experimental Neuropsychology*, *41*(10), 1015–1023. doi: [10.1080/13803395.2019.1644294](https://doi.org/10.1080/13803395.2019.1644294)
- Nijdam-Jones, A., & Rosenfeld, B. (2017). Cross-cultural feigning assessment: A systematic review of feigning instruments used with linguistically, ethnically, and culturally diverse samples. *Psychological Assessment*, *29*(11), 1321–1336. doi: [10.1037/pas0000438](https://doi.org/10.1037/pas0000438)
- Oudman, E., Krooshof, E., van Oort, R., Lloyd, B., Wijnia, J. W., & Postma, A. (2020). Effects of Korsakoff Amnesia on performance and symptom validity testing. *Applied Neuropsychology: Adult*, *27*(6), 549–557. doi: [10.1080/23279095.2019.1576180](https://doi.org/10.1080/23279095.2019.1576180)
- Paul, R., Rhee, G., Baker, L. M., Vaida, F., Cooley, S. A., & Ances, B. M. (2017). Effort and neuropsychological performance in HIV-infected individuals on stable combination antiretroviral therapy. *Journal of Neurovirology*, *23*(5), 725–733. doi: [10.1007/s13365-017-0557-5](https://doi.org/10.1007/s13365-017-0557-5)
- Ponsford, J., Downing, M., & Pechlivanidis, H. (2020). The impact of cultural background on outcome following traumatic brain injury. *Neuropsychological Rehabilitation*, *30*(1), 85–100. doi: [10.1080/09602011.2018.1453367](https://doi.org/10.1080/09602011.2018.1453367)

- Rai, J. K., & Erdodi, L. A. (2021). Impact of criterion measures on the classification accuracy of TOMM-1. *Applied Neuropsychology: Adult*, 28(2), 185–196. doi: [10.1080/23279095.2019.1613994](https://doi.org/10.1080/23279095.2019.1613994)
- Rees, L. M., Tombaugh, T. N., & Boulay, L. (2001). Depression and the test of memory malingering. *Archives of Clinical Neuropsychology*, 16(5), 501–506. [https://doi.org/10.1016/S0887-6177\(00\)00064-0](https://doi.org/10.1016/S0887-6177(00)00064-0)
- Rees, L. M., Tombaugh, T. N., Gansler, D. A., & Moczynski, N. P. (1998). Five validation experiments of the Test of Memory Malingering (TOMM). *Psychological Assessment*, 10(1), 10–20. doi: [10.1037/1040-3590.10.1.10](https://doi.org/10.1037/1040-3590.10.1.10)
- Salazar, X. F., Lu, P. H., Wen, J., & Boone, K. B. (2007). The use of effort tests in ethnic minorities and in non-English-speaking and English as a second language populations. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 405–427). New York, NY: The Guilford Press.
- Schroeder, R. W., Martin, P. K., Heinrichs, R. J., & Baade, L. E. (2019). Research methods in performance validity testing studies: Criterion grouping approach impacts study outcomes. *The Clinical Neuropsychologist*, 33(3), 466–477. doi: [10.1080/13854046.2018.1484517](https://doi.org/10.1080/13854046.2018.1484517)
- Sherer, M., Sander, A. M., Ponsford, J., Vos, L., Poritz, J. M. P., Ngan, E., & Leon Novelo, L. (2020). Patterns of cognitive test scores and symptom complaints in persons with TBI who failed performance validity testing. *Journal of the International Neuropsychological Society*, 26(9), 932–938. doi: [10.1017/S1355617720000351](https://doi.org/10.1017/S1355617720000351)
- Sherry, N., Ernst, N., French, J. E., Eagle, S., Collins, M., & Kontos, M. (2021). Performance validity testing in patients presenting to a specialty clinic with a mild traumatic brain injury. *Journal of Head Trauma Rehabilitation*, Advance online publication. doi: [10.1097/HTR.0000000000000692](https://doi.org/10.1097/HTR.0000000000000692)
- Sherman, E. M. S., Slick, D. J., & Iverson, G. L. (2020). Multidimensional malingering criteria for neuropsychological assessment: A 20-year update of the malingered neuropsychological dysfunction criteria. *Archives of Clinical Neuropsychology*, 35(6), 735–764. doi: [10.1093/arclin/acia019](https://doi.org/10.1093/arclin/acia019)
- Sieck, B. C., Smith, M. M., Duff, K., Paulsen, J. S., & Beglinger, L. J. (2012). Symptom validity test performance in the Huntington disease clinic. *Archives of Clinical Neuropsychology*, 28(2), 135–143. doi: [10.1093/arclin/acs109](https://doi.org/10.1093/arclin/acs109)
- Strutt, A. M., Scott, B. M., Lozano, V. J., Tieu, P. G., & Peery, S. (2012). Assessing sub-optimal performance with the Test of Memory Malingering in Spanish speaking patients with TBI. *Brain Injury*, 26(6), 853–863. doi: [10.3109/02699052.2012.655366](https://doi.org/10.3109/02699052.2012.655366)
- Suchy, Y. (2019). Introduction to special issue: Current trends in empirical examinations of performance and symptom validity. *The Clinical Neuropsychologist*, 33(8), 1349–1353. doi: [10.1080/13854046.2019.1672334](https://doi.org/10.1080/13854046.2019.1672334)
- Sullivan, B. K., May, K., & Galbally, L. (2007). Symptom exaggeration by college adults in attention-deficit hyperactivity disorder and learning disorder assessments. *Applied Neuropsychology*, 14(3), 189–207. doi: [10.1080/09084280701509083](https://doi.org/10.1080/09084280701509083)
- Sweet, J. J., Heilbronner, R. L., Morgan, J. E., Larrabee, G. J., Rohling, M. L., Boone, K. B., Kirkwood, M. W., Schroeder, R. W., Suhr, J. A., & Conference Participants. (2021). American Academy of Clinical Neuropsychology (AACN) 2021 consensus statement on validity assessment: Update of the 2009 AACN consensus conference statement on neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist*, 35(6), 1053–1106. doi: [10.1080/13854046.2021.1896036](https://doi.org/10.1080/13854046.2021.1896036)
- Sweet, J. J., Klipfel, K. M., Nelson, N. W., & Moberg, P. J. (2021). Professional practices, beliefs, and incomes of U.S. neuropsychologists: The AACN, NAN, SCN 2020 practice and “salary survey”. *The Clinical Neuropsychologist*, 35(1), 7–80. doi: [10.1080/13854046.2020.1849803](https://doi.org/10.1080/13854046.2020.1849803)
- Tombaugh, T. N. (1996). *Test of memory malingering: TOMM*. New York, NY: Multi-Health Systems.
- Uiterwijk, D., Wong, D., Stargatt, R., & Crowe, S. F. (2021). Performance and symptom validity testing in neuropsychological assessments in Australia: A survey of practises and beliefs. *Australian Psychologist*, 56(5), 355–371. doi: [10.1080/00050067.2021.1948797](https://doi.org/10.1080/00050067.2021.1948797)
- Victor, T. L., Boone, K. B., Serpa, J. G., Buehler, J., & Ziegler, E. A. (2009). Interpreting the meaning of multiple symptom validity test failure. *The Clinical Neuropsychologist*, 23(2), 297–313. doi: [10.1080/13854040802232682](https://doi.org/10.1080/13854040802232682)
- Walter, J., Morris, J., Swier-Vosnos, A., & Pliskin, N. (2014). Effects of severity of dementia on a symptom validity measure. *The Clinical Neuropsychologist*, 28(7), 1197–1208. doi: [10.1080/13854046.2014.960454](https://doi.org/10.1080/13854046.2014.960454)
- Wodushek, T. R., & Domen, C. H. (2020). Comparing two models of performance validity assessment in patients with Parkinson’s disease who are candidates for deep brain stimulation surgery. *Applied Neuropsychology: Adult*, 27(1), 9–21. doi: [10.1080/23279095.2018.1473251](https://doi.org/10.1080/23279095.2018.1473251)