

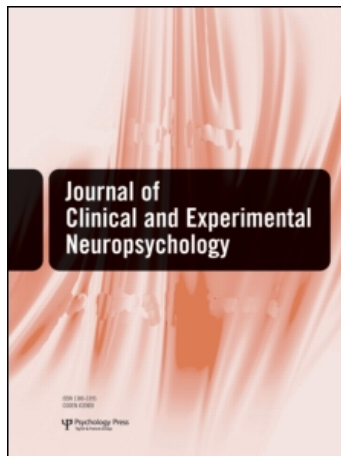
This article was downloaded by: [University of Kentucky]

On: 11 April 2011

Access details: Access Details: [subscription number 930576194]

Publisher Psychology Press

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Journal of Clinical and Experimental Neuropsychology

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713657736>

### New measures to detect malingered neurocognitive deficit: Applying reaction time and event-related potentials

Victoria L. Vagnini<sup>a</sup>; David T. R. Berry<sup>a</sup>; Jessica A. Clark<sup>a</sup>; Yang Jiang<sup>a</sup>

<sup>a</sup> University of Kentucky, Lexington, KY, USA

First published on: 26 February 2008

**To cite this Article** Vagnini, Victoria L. , Berry, David T. R. , Clark, Jessica A. and Jiang, Yang(2008) 'New measures to detect malingered neurocognitive deficit: Applying reaction time and event-related potentials', Journal of Clinical and Experimental Neuropsychology, 30: 7, 766 — 776, First published on: 26 February 2008 (iFirst)

**To link to this Article:** DOI: 10.1080/13803390701754746

**URL:** <http://dx.doi.org/10.1080/13803390701754746>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## New measures to detect malingered neurocognitive deficit: Applying reaction time and event-related potentials

Victoria L. Vagnini, David T. R. Berry, Jessica A. Clark, and Yang Jiang

University of Kentucky, Lexington, KY, USA

The ability of the Test of Memory Malingering (TOMM), reaction times (RTs), and event-related potentials (ERPs) to detect malingered neurocognitive deficit (MNCD) was examined in 32 normal individuals answering under honest (HON;  $n=16$ ) or malingering (MAL;  $n=16$ ) instructions as well as in 15 patients with traumatic brain injury (TBI) who answered under honest instructions. Overall, the TOMM was the most effective at classifying groups. However, new accuracy, RT, and ERP measures reached promising hit rates in the range of 71–88%. In particular, the difference in frontal versus posterior ERP obtained during an old–new task was effective at classifying MAL versus TBI (hit rate=87%).

**Keywords:** Electroencephalography; Event-related potential; Malingered neurocognitive deficit; Memory malingering; Traumatic brain injury.

In the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision* (American Psychiatric Association, 2000), malingering is defined as the deliberate fabrication or gross exaggeration of physical or psychological symptoms in pursuit of external goals such as avoiding duty or obtaining financial compensation. Malingering is distinguished from somatoform disorders in that in the latter, symptoms are not thought to be under voluntary control and are produced in pursuit of internal goals, such as resolving unconscious conflicts. Malingering is discriminated from factitious disorder in that in the latter, although symptoms are consciously produced, they appear to be in pursuit of internal goals such as achieving the sick role.

Malingered symptoms have been of increasing concern in neuropsychological assessment because most cognitive tests require optimal effort from the evaluatee in order to achieve results that accurately reflect brain functioning. In situations in which

patients are being evaluated for disability pensions or monetary compensation for damages sustained in accidents, the potential motivation to exaggerate or even fabricate problems is obvious. In fact, surveys of forensic neuropsychologists have suggested that up to 40% of individuals undergoing evaluations following mild head injuries may be malingering deficits (Mittenberg, Patton, Canyock, & Condit, 2002).

Neuropsychologists have responded to this concern by developing and validating a number of tests for the identification of inadequate effort during neuropsychological examinations. Vickery, Berry, Inman, Harris, and Orey (2001) published a meta-analysis of widely studied malingering tests and found that most had large effect sizes, high specificity (percentage of honest patients correctly classified), and moderate sensitivity (percentage of feigners correctly classified). However, recently, concerns have been raised that unscrupulous

The paper is based on V. L. Vagnini's PhD dissertation in partial fulfillment of the requirements for a doctoral degree in clinical psychology from the University of Kentucky. This research was supported in part by the American Forensic Academy for dissertation award funds, the Pilot fund of Behavioral Science Department at the University of Kentucky to VLV, and National Institutes of Health (NIH) Grant AG00986 to YJ.

Address correspondence to Yang Jiang, Department of Behavioral Science, University of Kentucky College of Medicine, Lexington, KY, USA (E-mail: yjiang@uky.edu).

parties to legal proceedings may be “coaching” plaintiffs to avoid being detected by commonly used malingering tests, and investigators have shown that internet search engines can easily identify information compromising these tests (Suhr & Gunstad, 2007).

In addition to validating new tests of malingered neurocognitive deficit (MNCD) such as the Test of Memory Malingering (TOMM; Tombaugh, 1996), other approaches have explored aspects of the evaluatee’s responses that are under less conscious control, such as reaction time (RT) and brain activity using electroencephalography (EEG). In general, these two approaches hypothesize that the physiological markers of neural processes that are normally “automatic” in cooperative test-takers are distorted in the malingering evaluatee. For example, because a malingerer may be consciously deciding to override a correct with an incorrect response, reaction times may be increased relative to the cooperative evaluatee who merely determines what the correct response is and gives it. Similarly, it is possible that EEG signals may be affected by additional cognitive processes engaged in by malingerers. Recently, Kozel and colleagues (2005) found differences in brain activity, using functional magnetic resonance imaging (fMRI) in the anterior cingulate, orbitofrontal cortex, and dorsolateral prefrontal cortex regions, that were able to detect deception in 90% of the participants. RT did not provide any additional detection ability; however there was a 3.5-s lag time before participants could respond, which may have obscured group differences due to lying.

Reaction time is easily obtained during testing on computer-administered procedures. Consistent with the earlier noted hypothesis that malingering evaluatees must devote extra time to selection of incorrect responses, some reports have documented slower RTs in feigners than in both normal and brain-injured control groups (Strauss, Spellacy, Hunter, & Berry, 1994; van Gorp et al., 1999). In contrast to these findings, Rose, Hall, and Szalda-Petree (1995) reported that honestly responding brain-injured patients had slower reaction times than did normals instructed to malingering. However, Rose et al. used a computerized version of the Portland Digit Recognition Test (PDRT) in which there were not instructions to respond as quickly as possible, which may account, in part, for their results. In addition to assessing mean RTs, the pattern of RTs has also been explored as a malingering detection strategy. Osimani, Alon, Berger, and Abarbanel (1997) attempted to detect individuals feigning brain injury using a commonly occurring cognitive phenomenon, the “Stroop

effect.” These authors found that honest normals and brain-injured patients both showed the Stroop effect, whereas malingerers (both uninformed and coached malingering groups) showed an inverted Stroop effect. Thus, examination of both mean and pattern of RTs has shown some initial promise for detecting malingering.

Event-related potentials (ERPs; i.e., averaged EEG signals) have also been explored for detecting malingering. ERPs are time locked to an event or stimulus, such as visual presentation of a picture. ERP signals have high temporal resolution and can reflect brain responses in milliseconds. ERP waveforms are labeled P for positive and N for negative deviations. For example, the P300 is a positive-going ERP waveform usually seen 300 to 800 milliseconds following stimulus presentation. P300, also known as P3, has been found to be an indicator of several important cognitive processes, such as attention and recognition memory. For example, Paller and Kutas (1992) showed that P300 has a larger amplitude when a subject is shown a previously presented stimulus (old), and a lower amplitude when a subject views a novel stimulus. Thus, a larger P300 waveform implies that a subject is, at some level, recognizing a previously shown stimulus, even if the overt response is to deny knowledge of the item. This ERP old–new effect has been shown to be reliable in memory tasks (Rugg, 1995, for a review). Interestingly, the old–new recognition task is very similar to standard malingering tests such as the TOMM, although perhaps slightly more difficult. Additionally, in the old–new task, RTs to old items are consistently faster than RTs to new items. Thus, ERP measures from the old–new task have the potential to provide an indirect test of whether a subject has any implicit memory of a previously presented stimulus, a characteristic of obvious possible application for detecting individuals falsely claiming amnesia. An additional potential measure from the old–new task arises from the possibility that the pattern of visual evoked potentials may also be altered by an attempt to malingering memory loss, such as different distributions of electrical activity across the scalp in honest versus feigning participants.

To date, several research publications have investigated the application of P300 and the pattern of visual evoked potentials to detection of malingered memory deficits, including reports by Ellwanger, Rosenfeld, Sweet, and Bhatt (1996), Ellwanger, Tenhula, Rosenfeld, and Sweet (1999), Rosenfeld et al. (1999), Rosenfeld et al. (1998), Rosenfeld, Sweet, Chuang, Ellwanger, and Song (1996), Tardif, Barry, Fox, and Johnstone (2000),

Tardif, Barry, and Johnstone (2002). Results from these studies were somewhat mixed and were limited in part by the relatively rare inclusion of a neurological patient group answering honestly. However, they do suggest some merit for the application of ERP to detection of feigned memory deficit and support additional work in the area.

The present study was undertaken to evaluate the potential utility of adding RT and ERP procedures to a well-validated test of MNCD: the Test of Memory Malingering. This investigation is one of the first to compare established MNCD procedures directly with newer physiological measures for the detection of feigned memory impairments. Additionally, it is one of very few physiological studies that include, in addition to normals answering honestly and normals asked to feign cognitive deficits, a group of moderately to severely head-injured patients answering under honest instructions.

## METHOD

### Participants

Three groups of participants were recruited: a group of normal volunteers with no history of head injury instructed to perform honestly (HON); a group of normal volunteers with no history of head injury instructed to malingering deficits (MAL); and a clinical group with documented brain injury instructed to perform honestly (TBI). A total of 16 right-handed participants were included initially in each of the three groups (48 total participants). Normal volunteers were recruited through fliers posted at University of Kentucky Introductory Psychology subject pool classes and were required to have no history of TBI. Participants in the TBI group were recruited through fliers at a local private practice neuropsychologist's office and an advertisement in a local newspaper. TBI participants were required to have sustained their brain injury at least 2 years prior to the study, be their own legal guardian, score  $\geq 24$  on a cognitive screening measure, the Mini Mental Status Exam (MMSE; Folstein, Folstein, & McHugh, 1975), agree to sign a release for their medical records to be reviewed, and to have documentation of an emergency room Glasgow Coma Scale (GCS)  $< 13$  following their injury. In addition, TBI participants were required to have been hospitalized after their injury, experienced a loss of consciousness (LOC)  $\geq 30$  min, and not be currently seeking compensation for their brain injury. After medical records were obtained, one TBI participant was

dropped from all analyses because the head injury would be classified as "mild" (GCS  $> 12$ ), leaving 15 participants in the TBI group. Medical records indicated that TBI participants had a mean emergency room GCS score of 8.7 ( $SD=2.9$ ), had a mean duration of loss of consciousness of 7.2 days ( $SD=12.0$ ), and were an average of 13 years postinjury ( $SD=7.2$ ), and that most (73.3%) were injured in moving-vehicle accidents. Computed tomography (CT) and magnetic resonance imaging (MRI) scans indicated brain injury in varied regions from brain stem, frontal, temporal, occipital, and parietal injuries in both the right and left hemispheres. All participants were compensated \$10/hour for their time. TBI participants were compensated an extra \$20 as a travel stipend as they were being recruited outside the greater Lexington area.

### Materials

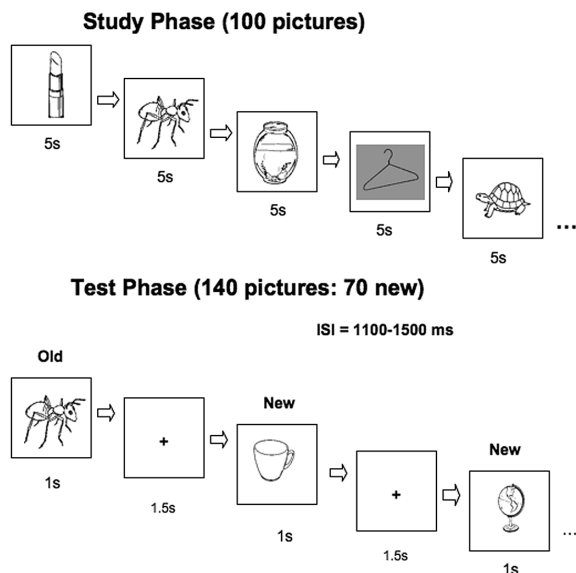
All participants were hooked up with a 32-channel electrode EEG cap using a Neuroscan system. Participants performed two tasks: a computerized version of the Test of Memory Malingering (TOMM-C) and an old-new memory recognition task. In addition, all participants completed a brief demographic questionnaire, a Beck-II (BDI-II) depression inventory, and a Wechsler Test of Adult Reading (WTAR).

#### *Test of Memory Malingering-computerized version (TOMM-C)*

For the computerized TOMM task, the standard computer format was used with line drawings presented on a computer screen. Stimulus pictures were displayed on a computer screen approximately 65 cm from the participants. Object size was approximately 8 cm by 10 cm, displayed in front of a white background with a black border. In order to protect the integrity of the TOMM, specific details about the nature of this test are not provided here.

#### *Old-new task*

The old-new task (see Figure 1) began with a study phase of 100 new line drawings (Cycowicz, Friedman, Rothstein, & Snodgrass, 1997; Snodgrass & Vanderwart, 1980). Stimulus pictures were displayed on a computer screen approximately 65 cm from the participants with a visual angle of approximately 7 degrees. Object size was approximately 8 cm by 6 cm, displayed in front of a black background. These 100 pictures were



**Figure 1.** An illustration of the study and test phases of the old-new memory recognition task.

presented for 5 seconds each during the study phase, and participants were instructed to memorize each picture. After a short break, all 100 pictures were studied again for the second time. After studying the 100 pictures a second time, the participants viewed 140 pictures, presented one at a time (70 old pictures from the study phase and 70 foils that had not been presented to the participant). For each picture, the participant decided whether the drawing was “new” or “old” and clicked a corresponding key on the keyboard. Stimulus onset was delayed randomly at 100 ms, 300 ms, or 500 ms in order to avoid an expectation effect of exactly when the pictures would appear on the screen. Each picture was presented for 1,000 ms with an interstimulus interval (ISI) of 1,100–1,500 ms. Fixations were presented for 1,500 ms. Participants stared at the screen without blinking until a fixation cross (+) appeared, indicating that they could blink. The old-new task took approximately 9.5 minutes.

## ERP recording

The electroencephalogram (EEG) was recorded from 32 scalp sites using an electrode cap with silver/silver chloride inserts. All scalp electrodes were referenced to an electrode placed between CZ and CPZ. In order to detect eye blink artifacts that would distort the ERP data measured at the scalp, horizontal electrooculogram (HEOG) and vertical electrooculogram (VEOG) were recorded with two pairs of electrodes (one pair placed between the

eyebrows and the other below the left eye, and another pair placed on the outer edge of each eye). EEG was recorded continuously during the tasks. EEG signals were filtered with a bandpass of 0.1 to ~40 Hz and were sampled at a rate of 500 Hz. Electrode impedances did not exceed 10 k $\Omega$ , ensuring a good connection between the electrode and scalp for an accurate ERP recording. Averaged ERPs were formed offline from correct-response trials and incorrect-response trials, and all responses for the old-new task were screened to be free of ocular and movement artifacts (more than  $\pm 75$  mV). For the old-new task each scalp site resulted in two separate ERP waveforms (old and new). The mean numbers of individual trials per waveform that were free from movement or eye blink artifacts that distort the ERP recording were calculated.

The mean number of correct items averaged for each HON participant ranged from 46 to 60 for the old-new task. There were far fewer trials averaged for the incorrect responses for HON ranging from 8 to 14. While this number is well below the ideal number of trials for an average wave, it allowed for cautious comparison with MAL and TBI that had higher error rates. Examining the incorrect responses is warranted here despite the reduced number of trials because it may reveal a difference related to malingering that is obscured when comparing correct and all responses. The mean number of correct and all items averaged for each TBI participant ranged from 30 to 51 for the old-new effect and 18 to 28 for incorrect responses. The mean number of correct and all items averaged for each MAL participant ranged from 27 to 51, and the average number of trials for incorrect responses ranged from 20 to 21.

## Procedure

Participants completed demographic, BDI-II, and WTAR questionnaires/tests under standard instructions when they arrived at the lab. Participants with a history of brain injury also filled out a release of information for obtaining data about their brain injury from their medical records. A standard eye exam was administered to all participants in order to make sure their vision was 20/20 or corrected to 20/20. Participants in the non-brain-injured and malingering groups were selected to match the age, education, and gender of participants in the TBI group as much as possible. Prior to beginning the TOMM-C, participants were given one of two sets of instructions. Honest and brain-injured groups received instructions to give

their best effort on the tasks, and participants in the malingering group were instructed to pretend that they had memory problems as a result of an accident. After reading the instruction set aloud, the RA asked each participant to paraphrase his or her understanding of how to respond to the tasks and did not move to the first task until the participant was clear about how to respond. Next, participants were asked to read over the instructions for the TOMM-C, and two sample questions were administered to ensure that the participant knew the instructions for the task. All participants were administered Trial 1 and Trial 2 of the TOMM test on the computer in a sound-proof testing chamber. During the standard 15-minute delay between Trial 2 and the retention trial (T3) of the TOMM, two research assistants (RAs) began to put a 32-channel EEG cap on the participant.

After completing the retention trial of the TOMM, the RAs finished putting on the EEG cap (approximately 10–15 minutes). After the cap was secure, the participants read the instructions for the old–new task, and the RA answered any questions. Participants then studied 100 new line drawings that were different from the TOMM pictures on a computer screen presented for 5 seconds each. The RA watched the participant take a brief practice trial (approximately 3 minutes) of the old–new task to ensure that the participant understood the instructions for the task and the instructions regarding blinking control. The participant then performed the old–new task with EEG recording. It must be noted that the RAs were not blind to the condition for HON and MAL; it was not deemed necessary for this study where the three tasks to be compared were computer administered and not subject to any bias effects by the RA.

### Statistical analysis

Mean response times (RTs), correct responses (accuracy), and ERPs in three groups were averaged for each condition. Unless otherwise noted, all analyses were conducted on data from 16 participants in the HON group, 16 in the MAL group, and 15 in the TBI group (47 participants total). The number of participants changed slightly (1–2 fewer) for some of the ERP analyses due to occasional “bad” channels, which led to a participant being dropped from an analysis. The change is reflected in the degrees of freedom in the reported *F*-values. The ERP responses in the old–new task in the three groups were compared for the six midline electrodes in a 3 (group)  $\times$  2 (novelty)  $\times$  6 (electrode) mixed design repeated measures

analysis of variance (ANOVA) for the three types of response (correct, incorrect, and all responses). The time intervals examined for ERPs in the old–new task were 0–100 ms, 100–250 ms, 250–400 ms, 400–650 ms, 650–750 ms, and 750–900 ms. In the case of significant interactions between group and experimental effects, main effects of noninterest might not be reported. For all analyses, the Greenhouse–Geisser correction was used to identify significant interaction effects with the group variable, and post hoc pairwise comparisons were conducted using Tukey’s HSD test.

## RESULTS

There were no significant differences between groups (all *ps* > .05) for any demographic variables, including age, education, IQ estimate (WTAR), acute emotional state (BDI-II), gender, race, and marital status, as shown in Table 1, based on ANOVAs or  $\chi^2$  as appropriate.

### Accuracy results

To determine whether the three study groups obtained different accuracy rates, an initial multivariate analysis of variance (MANOVA) was performed on the TOMM-C and old–new task as presented in Table 2. Results indicated a significant overall effect: Wilks’s  $\Lambda$  = .144, *F*(8, 82) = 16.77, *p* < .001. Based on ANOVA, there was a significant difference between groups for TOMM Trial 2, *F*(2, 44) = 72.43, *p* < .0001, as well as TOMM Trial 3, *F*(2, 44) = 77.43, *p* < .0001. Post hoc comparisons

**TABLE 1**  
Demographic information for the three study groups

Variable	Group		
	HON	MAL	TBI
Age	36.2 (12.2)	32.7 (12.8)	40.5 (11.7)
Education	15.4 (2.3)	15.7 (2.5)	14.3 (1.9)
WTAR	110.3 (11.3)	111.8 (7.8)	105.5 (6.9)
BDI-II	5.1 (7.6)	2.5 (4.2)	6.9 (6.3)
Gender <sup>a</sup>	56.3	62.5	46.7
Race <sup>b</sup>	68.8	93.8	100.0
Marital status <sup>c</sup>	50.0	56.3	53.3
<i>N</i>	16	16	15

*Note.* Means; standard deviations in parentheses. HON=honest responders; MAL=malingering responders; TBI=traumatic brain injury group; WTAR=Wechsler Test of Adult Reading; BDI-II=Beck Depression Inventory–II. Age and education given in years. No variables were significantly different (*ps* > .05).

<sup>a</sup>Percentage female. <sup>b</sup>Percentage white. <sup>c</sup>Percentage single.

**TABLE 2**  
Accuracy results for the three study groups

Test	Variable	Group			MAL vs. TBI ( <i>d</i> )
		HON	MAL	TBI	
TOMM-C	T2***	99.9 <sup>a</sup> (0.5)	64.0 <sup>b</sup> (16.7)	99.6 <sup>a</sup> (1.5)	3.0
	T3***	100 <sup>a</sup> (0.0)	61.1 <sup>b</sup> (17.7)	99.8 <sup>a</sup> (0.7)	3.0
Old–new task	Old items***	93.5 <sup>a</sup> (5.6)	64.0 <sup>b</sup> (17.6)	81.4 <sup>a</sup> (18.5)	1.0
	New items**	87.1 <sup>a</sup> (9.8)	67.2 <sup>b</sup> (14.5)	74.5 <sup>b</sup> (18.3)	0.4
	All items***	90.3 <sup>a</sup> (5.8)	65.6 <sup>b</sup> (12.2)	75.2 <sup>b</sup> (17.0)	0.7
	Diff. score*	4.5 <sup>ab</sup> (7.7)	–2.3 <sup>a</sup> (14.8)	8.7 <sup>b</sup> (9.9)	0.9

*Note.* Mean percentages correct; standard deviations in parentheses. HON=honest responders; MAL=malingering responders; TBI=traumatic brain injury group. TOMM-C=Test of Memory Malingering–computerized. T2= Trial 2. T3= Trial 3. Diff. score=old – new correct items from the old–new task. Cohen's *d* values are absolute values of mean differences between MAL and TBI for variables that were significantly different since this comparison is the most clinically relevant.

<sup>ab</sup> denotes groups significantly different from each other; same letter means groups are not different from each other.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

indicated that for both trials MAL had significantly lower accuracy (61–64%) than HON and TBI, which both had almost perfect accuracy ( $p < .0001$  for both comparisons). These findings indicate that the accuracy scores from this standard malingering test were successful in documenting differences between MAL and both honest responding groups (HON and TBI). Table 2 also presents the effect size (ES) data in Cohen's *d* (Cohen, 1998) between MAL and TBI (since it is the most clinically relevant comparison for the purposes of this study) for accuracy variables. For TOMM-C accuracy variables the ES for the difference between HON and TBI compared to MAL was very large (Cohen's  $d=3.1$  and  $3.0$ , respectively) and further supported the strength of this traditional method of detecting MNCD.

Table 2 also presents accuracy data from the old–new task. As may be seen here, there were statistically significant effects of group on all old–new variables. However, post hoc comparisons indicated that only the old items and the difference score between old and new items from the task had the same general pattern as that seen for the TOMM variables (HON=TBI > MAL), with all remaining old–new variables showing a less promising pattern for discriminating the TBI from MAL group. The *d*-scores for the TBI versus MAL comparisons were generally much weaker than for the TOMM, although the old items from the old–new task were fairly robust at 1.0.

## RT results

Table 3 presents RT results from the groups on the TOMM-C and the old–new task. For the TOMM-C,

a MANOVA indicated a significant overall effect: Wilks's  $\Lambda=.656$ ,  $F(8, 82)=2.4$ ,  $p < .05$ . RTs for correct and all responses in Trial 3 as well as an overall RT (all responses for Trial 2 and Trial 3) differed significantly across the three groups. However, for the two significant RT variables, follow-up analyses indicated that the only significant pairwise comparisons were between HON and MAL. Thus, RT results from the TOMM-C do not appear promising for discrimination of MAL and TBI groups, as reinforced by the modest *d*-values.

Table 3 also presents RT results from the old–new task. In general, the TBI group had the longest RTs. Further inspection of these results indicates that there were almost no significant differences between the MAL and TBI groups, with the exception of reaction time from all responses to new items. Unfortunately, in this case, the TBI group was significantly slower than the MAL group. Overall, results in Table 3 suggest that prolonged response latencies on the TOMM-C and old–new tests are not viable markers of feigned deficits.

## Event-related potential (ERP) results

### ERP mean amplitude for midline electrodes

Visual inspection of ERP averaged waveforms was conducted for the three groups to determine time intervals that showed the greatest difference among MAL, HON, and TBI. A significant interaction between Novelty×Group,  $F(2, 44)=3.3$ ,  $p < .05$ , was found across the midline electrodes for all responses at the 400–650-ms interval. This “old–new effect” is a classic ERP finding reported in honest and normal healthy participants, where ERP P300 responses to studied items are typically

**TABLE 3**  
Reaction time (in ms) results for the three study groups

			Group			
Test	Variable		HON	MAL	TBI	MAL vs. TBI ( <i>d</i> )
TOMM-C	T3	Correct responses	1,084 (266)	2,314 (2286)	1,394 (418)	0.6
		All responses*	1,084 <sup>a</sup> (266)	1,632 <sup>b</sup> (635)	1,395 <sup>ab</sup> (417)	0.4
		Overall: T2 & T3 all responses*	1,152 <sup>a</sup> (341)	1,592 <sup>b</sup> (585)	1,418 <sup>ab</sup> (356)	0.4
Old–new task	Old	Correct responses	732 (63)	797 (91)	844 (91)	
		Incorrect responses	825 (167)	791 (268)	880 (266)	
		All responses*	732 <sup>ac</sup> (67)	797 <sup>ab</sup> (93)	855 <sup>bc</sup> (93)	0.6
	New	Correct responses	793 (78)	823 (120)	937 (93)	
		Incorrect responses	837 (173)	781 (269)	862 (154)	
		All responses*	798 <sup>ad</sup> (80)	811 <sup>a</sup> (119)	921 <sup>bd</sup> (84)	1.1
	Difference (correct) new–old*		66 <sup>a</sup> (59)	14 <sup>ab</sup> (54)	66 <sup>b</sup> (84)	0.7

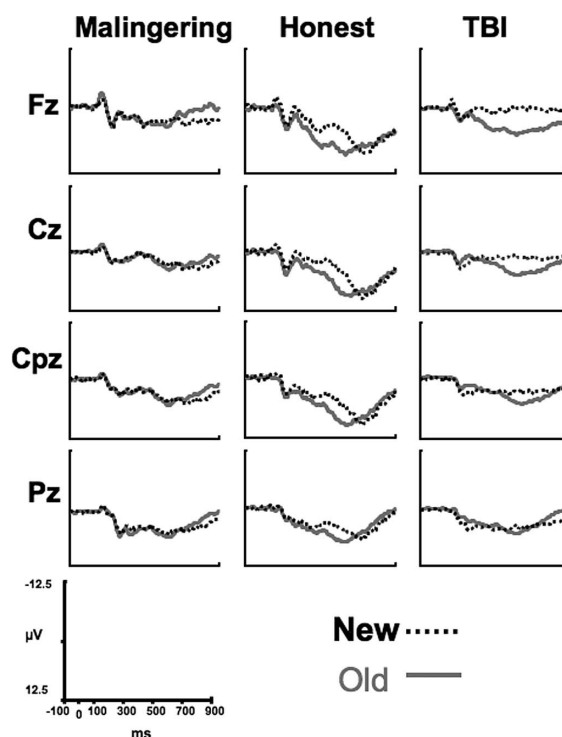
*Note.* Mean reaction times, in ms; standard deviations in parentheses. HON=honest responders; MAL=malingering responders; TBI=traumatic brain injury group. TOMM-C=Test of Memory Malingering–computerized. T2=Trial 2. T3=Trial 3. Cohen's *d* values are absolute values of mean differences between MAL and TBI for variables that were significantly different since this comparison is the most clinically relevant.

<sup>ab</sup>denotes groups significantly different from each other; same letter means groups are not different from each other; <sup>cd</sup>denotes within-group differences in reaction time between new and old items. \**p* ≤ .05.

larger and more positive going than those to new items in posterior electrodes. The mean ERP amplitude values for each group suggested that HON and TBI showed the typical ERP old–new effect. In contrast, MAL did not show much difference between old and new items in ERP responses. For HON, there was a significant difference in ERP mean amplitude such that studied items evoked stronger P300 ( $M=4.9$ ,  $SD=3.1$ ) than new items ( $M=2.9$ ,  $SD=2.2$ ). To a lesser extent, TBI followed the same typical trend of the old–new effect (for old items  $M=3.2$ ,  $SD=2.7$ ; for new items  $M=2.1$ ,  $SD=3.0$ ), even though the difference was not statistically significant ( $p=.24$ ). MAL, however, had no difference in the late positive ERP responses between old ( $M=2.3$ ,  $SD=3.2$ ) and new items ( $M=2.2$ ,  $SD=2.7$ ),  $p > .80$ . Figure 2 shows the ERP waves for the midline electrodes of the three groups for the old–new task using all responses, and the lack of an old–new effect in MAL is clearly visible. This finding suggests that the old–new effect is intact for HON, reduced but trending towards significant in TBI, and absent in MAL. Thus, these malingerers did not show the old–new effect, suggesting that feigning may eliminate or reduce the typical ERP responses for new versus studied items.

### Frontal versus posterior electrodes analysis

The differences between ERPs for frontal versus posterior electrodes were also examined (using the same time intervals mentioned above for the



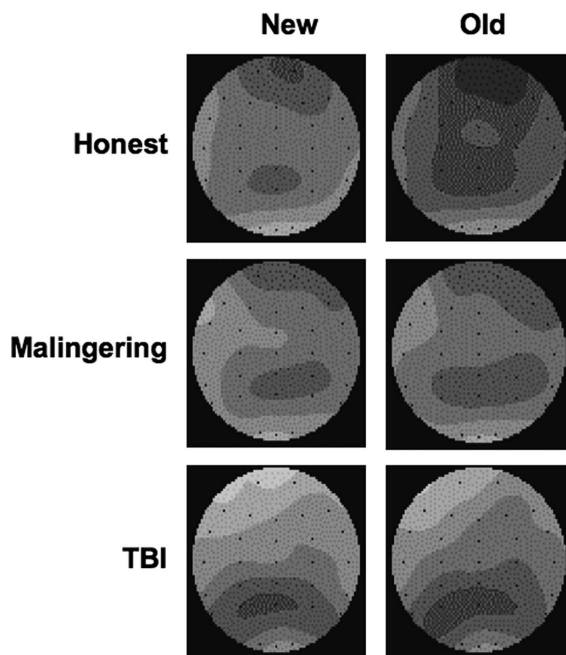
**Figure 2.** Average event-related potential (ERP) responses for the old–new task for honest, malingering, and traumatic brain injury (TBI) groups—frontal and central midline electrodes (all responses). The ERP old–new difference, seen in both honest and brain-injured groups, is visibly absent among malingerers.

midline analysis). The number of participants included in these analyses declined to 15 for the HON and MAL groups and 14 for TBI due to exclusion of some participants from the analysis



secondary to technical problems with their recordings. A significant group difference occurred in the frontal electrodes. For all three response types (correct, incorrect, and all), MAL had larger mean amplitudes than TBI and a moderate ES difference (mean Cohen's  $d=0.4$ ). In addition, HON had larger ERP mean amplitudes than MAL and a large ES difference (mean Cohen's  $d=0.9$ ). When healthy individuals malingered, the frontal electrodes show reduced ERP mean amplitude. Malingering appeared to diminish mean amplitude in frontal ERP responses, but this type of response was still larger than the ERPs for individuals with brain injury.

The most interesting finding, when comparing the mean ERP amplitude of frontal versus posterior electrodes, was the within-group differences that occurred during the intervals of 400–900 ms, which is typically P300 or late positive component. Figure 3 represents the frontal versus posterior electrode differences in recorded electrical activity among the groups using a topographical map across the scalp during the interval 400–650 ms. Figure 3 shows that when comparing activity from frontal and posterior electrodes within each group, HON had the strongest activity in the frontal area, TBI had the strongest responses in the posterior area, and MAL showed no significant difference between frontal and posterior activity (MAL frontal  $M=2.4$ ,  $SD=0.4$ ; posterior  $M=2.2$ ,  $SD=0.9$ ).



**Figure 3.** Scalp topography for the old–new task (400–650 ms) using all responses of honest, malingering, and traumatic brain injury (TBI) groups.

HON had larger mean ERP amplitude for frontal electrodes ( $M=4.6$ ,  $SD=1.0$ ) than for posterior electrodes ( $M=2.5$ ,  $SD=1.5$ ). Finally, individuals with TBI responding honestly used more posterior resources ( $M=3.0$ ,  $SD=1.2$ ) than frontal ( $M=0.9$ ,  $SD=0.6$ ). The pattern of activity (frontal vs. posterior) for the old–new task is important to note, because the topographic analysis for all three groups look different. The frontal–posterior difference might be an effective indicator to identify malingerers. These findings suggest that non-brain-injured individuals responding honestly show more frontal than posterior activity; however, when a person is malingered the frontal responses are reduced (less positive ERP amplitude). In addition, people with TBI have greater activity in posterior than in frontal electrodes. Using a difference score may help identify malingerers, because malingerers would be expected to have the smallest absolute value (close to “0”).

#### Classification statistics for accuracy, RT, and ERP data

To explore the relative accuracy of these measures for classifying individual results, the accuracy, RT, and ERP variables that generated the largest effect sizes for the MAL versus TBI contrast were selected. The established malingering test (TOMM-C), had perfect classification rates. To identify promising new variables, Table 4 presents selected accuracy, RT, and old–new variables and the cutting scores (using a 50% base rate of feigning), which obtained the highest overall hit rate for each variable in the most clinically relevant comparison, MAL versus TBI.

Results in Table 4 show that the maximal hit rate for new variables was achieved by the frontal–posterior ERP wave difference for incorrect responses (86.5%). Additionally, the new–old ERP wave difference was promising (77.1%), as was, to a lesser extent, accuracy for old items in the old–new task (74.4%). These three variables may merit additional research as malingering detectors.

#### DISCUSSION

The aim of this study was to explore the effectiveness of two new methods for detecting malingered neurocognitive deficit (MNCD)—that is, RT and ERP measures. These new measures were compared to an established method: the computerized version of the TOMM (TOMM-C). The TOMM-C accuracy results were consistent with prior research supporting its use as a malingering test (Rees,

**TABLE 4**  
Classification statistics for new accuracy, RT, and ERP variables comparing TBI versus MAL

	<i>Variables</i>	<i>Sensitivity</i>	<i>Specificity</i>	<i>Hit rate</i>	<i>NPP</i>	<i>PPP</i>
Accuracy	O/N task: Old item (<72% correct)	62.5	80.0	74.4	71.9	77.5
	O-N dif (≤6=MAL)	62.5	73.3	67.9	66.2	70.1
RT	TOMM-C: T3 RT (crct) (>1,575 ms)	62.5	80.0	71.3	68.1	75.8
	O/N task: O-N RT dif (<-76 ms=MAL)	81.3	60.0	70.7	76.2	67.0
ERP	N-O wave dif (<-75=MAL)	87.5	66.7	77.1	84.2	72.4
	F-P wave dif-wrong (<2.0=MAL)	80.0	92.9	86.5	82.3	91.9

*Note.* Malingering: base rate for simulation malingering=50%. MAL=malingering responders; TBI=traumatic brain injury group (clinical); TOMM-C=Test of Memory Malingering-computerized; T3= Trial 3; RT=reaction time. ERP=event-related potential. O/N task=old-new task; O-N=old - new items; ms=milliseconds; NPP=negative predictive power; PPP=positive predictive power. F-P=frontal-posterior.

Tombaugh, Gansler, & Moczynski, 1998; Tombaugh, 1997) and were able to predict group membership with complete accuracy. Despite the clear success of the TOMM-C in the present study, the likelihood of people evading detection likely increases as more information on existing tests becomes available on internet sites and in academic journals. Therefore, developing new measures and examining additional detection methods such as RT and ERP are important to allow continued success in identifying individuals who feign neuro-cognitive deficit.

Although mean RT results were not very successful in discriminating honest and feigning participants in the present study, the pattern of RTs showed some promise as a detection strategy. Thus, when comparing RTs to old versus new items in the old-new task, some power to differentiate honest and feigning participants appeared. Both HON and TBI produced the classic "old-new effect" with slower RTs to new pictures than to old pictures, whereas the MAL group failed to show such a difference. These results replicate prior research findings that there may be some validity in examining tasks with different RT patterns for different items to detect feigned cognitive impairment (Osmani et al., 1997; van Gorp et al., 1999). The old-new RT difference score had a success rate that was much better than simply comparing mean RTs, which suggests that this variable has some potential to detect MNCD. In a future study, examining an old-new RT difference when participants have a longer time to respond may further differentiate honest and malingering responders.

The ERP results across tasks in this study indicated that the ERP mean amplitude was largest for honestly responding healthy adults, and furthermore the effect of malingering in non-brain-injured participants appears to reduce ERP mean amplitude such that these responders look similar to participants with TBI (the mean ERP amplitude for

MAL was only slightly larger than that for TBI). These results are similar to prior research by Ellwanger et al. (1996) and contrary to findings of Ellwanger et al. (1999), which indicated that malingerers had a larger P300 than honest responders. While the present study examined multiple frontal and posterior electrodes, the prior studies predominantly examined the P300 component at electrode Pz only, which may account for some of the discrepancies.

The most salient ERP finding was that the malingerers did not show the classic ERP "old-new effect," which was in contrast to the honest responders (Figure 2). Honest responders' RTs (HON and TBI) differed between old and new items but this differentiation was absent in malingerers. Thus, malingering appears to reduce inter-item ERP differences. Similar to the finding from the frontal versus posterior analyses, malingering reduced ERP differences seen in the two honest responding groups. The P300 was the predominant component examined in prior research using ERP to detect MNCD. For the old-new task, HON had the classic "old-new effect" response (mean amplitude for old items > than that for new). Although the difference for TBI did not reach significance, the trend was in the right direction. Once again malingerers showed no difference between old and new items. In a review of existing studies in this area, Rosenfeld and Ellwanger (1999) theorized that the lack of difference in ERPs for different items in malingerers may be due to the increased workload that occurs when test-takers manipulate their responses (e.g., deliberately select incorrect answers). In the present study, malingerers did not show variation in ERPs to items in the old-new task that were seen in both intact and TBI responders performing honestly.

ERPs from the front half of the scalp revealed a significant difference among the groups for the old-new task. HON had a larger ERP wave

segment in frontal than in posterior electrodes, while TBI had larger mean amplitude in the posterior than in the frontal electrodes. There was no significant difference in ERPs of frontal versus posterior electrodes for MAL. These results are consistent with the prior reports that across three midline electrodes (Fz, Cz, and Pz) malingerers did not show differences found in an honest responding group (Rosenfeld et al., 1999). The distribution of activity in the non-brain-injured, honestly responding person appears to occur primarily in the frontal region, and in TBI the primary activity is in the posterior part of the brain (perhaps compensation for damage as a result of a head injury). The lack of difference for MAL may be due to a focus on manipulating responses that may reduce the activity in the frontal electrodes. This can be seen in comparisons indicating that HON had significantly larger mean amplitude in the frontal electrodes than did MAL. The reduced mean ERP amplitude in MAL may be due to several issues. Kok (2001) discussed several factors that affect P3 amplitudes, which may illustrate what led to reduced mean amplitude in the frontal electrodes in the malingering group. Reduced ERP amplitude may have been related to decreased attention to the primary task due to a secondary task of malingering. Task difficulty also leads to a decreased P3, and for MAL the additional demand of malingering may have increased the complexity of the task. Finally, reduced confidence in one's performance is related to a decrease in P3. Malingerers may not be sure if they are answering in a way that will appear as if they have brain damage, and their confidence thus may suffer compared to HON.

It should be mentioned that there have been a few studies that have examined moderate-to-severe TBI using ERPs. The findings generally indicate that groups with brain impairment have brain activity suggesting compensatory strategies compared to honest controls. Individuals with TBI generally have reductions in ERP mean amplitude and longer peak latencies in auditory oddball tasks (Polo, Newton, Rogers, Escera, & Butler, 2002; Rugg et al., 1993). In other words, prior ERP research with TBI supports the notion that there are group differences related to brain injury, but that these changes are due to compensation rather than a lack of conscious processing that would limit the implications that could be drawn in malingering research.

One limitation of this study is the lack of ability to correlate the area of brain injury with specific ERP results (such as frontal vs. posterior differences). The TBI participants had documented brain injury, but the injuries were widespread

across the brain. This may have limited confidence in any statements that might be made about how malingerers perform differently from TBI on the memory task used in this study. A future study examining participants with acute damage in a specific area of the brain may be able to provide further insights. Similarly, test scores of memory function were not available in the medical records obtained, and no memory measure was administered to the participants to document memory impairment in the TBI group. A future study including a memory test could document the memory function of the TBI and healthy volunteer groups. In addition, future studies attempting to use RT and/or ERP methods to detect MNCD should provide malingering participants with incentives for successful malingering to reflect real-world circumstances (Rogers, 1997). Another issue here is that the interval to respond for the old-new task was also very short (approximately 1.5 s), and this may have reduced the difference between the MAL and TBI groups. When forced to respond very quickly, MAL participants were able to adjust and responded faster than participants with cognitive impairment.

In conclusion, this is the first study that compared three relevant groups in a malingering paradigm using multiple EEG electrodes. Accuracy, RT, and ERP variables from a standard malingering test and two new cognitive tasks were examined to see how well each measure detected MNCD. The standard malingering test (TOMM) replicated prior data as the gold standard to detect MNCD. The newer RT pattern measures obtained from a TOMM-like task with no set response time appear to have reasonable success in detecting MNCD. Finally, certain ERP measures also looked promising for detection of feigning. Further research is needed to replicate the present results as well as to explore the generalizability of these findings to other settings and participants.

Original manuscript received 26 August 2007

Revised manuscript accepted 9 October 2007

First published online 26 February 2008

## REFERENCES

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- Cohen, J. (1998). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.
- Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture naming by young children: Norms for name agreement, familiarity, and

- visual complexity. *Journal of Experimental Child Psychology*, 65, 171–237.
- Ellwanger, J., Rosenfeld, J. P., Sweet, J. J., & Bhatt, M. (1996). Detecting simulated amnesia for autobiographical and recently learned information using the P300 event-related potential. *International Journal of Psychophysiology*, 23, 9–23.
- Ellwanger, J., Tenhula, W. N., Rosenfeld, J. P., & Sweet, J. J. (1999). Identifying simulators of cognitive deficit through combined use of neuropsychological test performance and event-related potentials. *Journal of Clinical and Experimental Neuropsychology*, 21, 866–879.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189–198.
- Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, 38, 557–577.
- Kozel, F. A., Johnson, K. A., Mu, Q., Grenesko, E. L., Laken, S. J., & George, M. S. (2005). Detecting deception using functional magnetic resonance imaging. *Biological Psychiatry*, 58, 605–613.
- Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology*, 24, 1094–1102.
- Osimani, A., Alon, A., Berger, A., & Abarbanel, J. M. (1997). Use of the Stroop phenomenon as a diagnostic tool for malingering. *Journal of Neurology, Neurosurgery, & Psychiatry*, 62, 617–621.
- Paller, K. A., & Kutas, M. (1992). Brain potentials during memory retrieval provide neurophysiological support for the distinction between conscious recollection and priming. *Journal of Cognitive Neuroscience*, 4, 375–391.
- Polo, M. D., Newton, P., Rogers, D., Escera, C., & Butler, S. (2002). ERPs and behavioural indices of long-term preattentive and attention deficits after closed head injury. *Neuropsychologia*, 40, 2350–2359.
- 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, 10–20.
- Rogers, R. (1997). *Clinical assessment of malingering and deception* (2nd ed.). New York: Guilford Press.
- Rose, F. E., Hall, S., & Szalda-Petree, A. D. (1995). Portland Digit Recognition Test–Computerized: Measuring response latency improves the detection of malingering. *The Clinical Neuropsychologist*, 9, 124–134.
- Rosenfeld, J. P., & Ellwanger, J. (1999). Cognitive psychophysiology in detection of malingered cognitive deficit. In J. J. Sweet (Ed.), *Forensic neuropsychology: Fundamentals and practice* (pp. 287–311). Lisse, The Netherlands: Swets & Zeitlinger.
- Rosenfeld, J. P., Ellwanger, J. W., Nolan, K., Wu, S., Bermann, R. G., & Sweet, J. (1999). P300 scalp amplitude distribution as an index of deception in a simulated cognitive deficit model. *International Journal of Psychophysiology*, 33, 3–19.
- Rosenfeld, J. P., Reinhart, A. M., Bhatt, M., Ellwanger, J., Gora, K., Sekera, M., et al. (1998). P300 correlates of simulated amnesia on a matching-to-sample task: Topographic analyses of deception vs. truth-telling responses. *International Journal of Psychophysiology*, 28, 233–248.
- Rosenfeld, J. P., Sweet, J. J., Chuang, J., Ellwanger, J., & Song, L. (1996). Detection of simulated malingering using forced choice recognition enhanced with event-related potential recording. *The Clinical Neuropsychologist*, 10, 163–179.
- Rugg, M. D. (1995). Memory and consciousness: A selective review of issues and data. *Neuropsychologia*, 33, 1131–1141.
- Rugg, M. D., Pickles, C. D., Potter, D. D., Doyle, M. C., Pentland, B., & Roberts, R. C. (1993). Cognitive brain potentials in a three-stimulus oddball task after closed head injury. *Neuropsychologia*, 31, 373–393.
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology and Human Learning*, 6, 174–215.
- Strauss, E., Spellacy, F., Hunter, M., & Berry, T. (1994). Assessing believable deficits on measures of attention and information processing. *Archives of Clinical Neuropsychology*, 9, 483–490.
- Suhr, J. A., & Gunstad, J. (2007). Coaching and malingering: A review. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 287–311). New York: Oxford Press.
- Tardif, H. P., Barry, R. J., Fox, A. M., & Johnstone, S. J. (2000). Detection of feigned recognition memory impairment using the old new effect of the event-related potential. *International Journal of Psychophysiology*, 36, 1–9.
- Tardif, H. P., Barry, R. J., & Johnstone, S. J. (2002). Event-related potentials reveal processing differences in honest vs. malingered memory performance. *International Journal of Psychophysiology*, 46, 147–158.
- Tombaugh, T. N. (1996). *The Test of Memory Malingering (TOMM)*. New York: Multi-Health Systems.
- Tombaugh, T. N. (1997). The Test of Memory Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment*, 9, 260–268.
- van Gorp, W. G., Humphrey, L. A., Kalechstein, A., Brumm, V. L., McMullen, W. J., Stoddard, M., et al. (1999). How well do standard clinical neuropsychological tests identify malingering? A preliminary analysis. *Journal of Clinical and Experimental Neuropsychology*, 21, 245–250.
- Vickery, C. D., Berry, D. T. R., Inman, T. H., Harris, M. J., & Orey, S. A. (2001). Detection of inadequate effort on neuropsychological testing: A meta-analytic review of selected procedures. *Archives of Clinical Neuropsychology*, 16, 45–73.