Portland Digit Recognition Test the Test of Memory Malingering and standard neuropsychological measures: Comparing efficacy the effects of coaching and face validity

Melody L. Huskey

The University of Montana

Let us know how access to this document benefits you.
Follow this and additional works at: https://scholarworks.umt.edu/etd

Recommended Citation
Huskey, Melody L., "Portland Digit Recognition Test the Test of Memory Malingering and standard neuropsychological measures: Comparing efficacy the effects of coaching and face validity" (2002). Graduate Student Theses, Dissertations, & Professional Papers. 4993. https://scholarworks.umt.edu/etd/4993

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.
Permission is granted by the author to reproduce this material in its entirety, provided that this material is used for scholarly purposes and is properly cited in published works and reports.

**Please check "Yes" or "No" and provide signature**

Yes, I grant permission ☑

No, I do not grant permission

Author's Signature: 

Date: 5/14/02

Any copying for commercial purposes or financial gain may be undertaken only with the author's explicit consent.

by

Melody L. Huskey

B.S., Boise State University, 1994

Presented in partial fulfillment of the requirements for the degree of Master of Arts

The University of Montana

May 2002

Approved by:

[Signatures]

Dean, Graduate School

5-14-02

Date
The Portland Digit Recognition Test, the Test of Memory Malingering, and Standard Neuropsychological Measures: Comparing Efficacy, the Effects of Coaching, and Face Validity

Director: Stuart Hall, Ph.D.

This research examined the performance of control participants and individuals feigning brain damage on the Test of Memory Malingering (TOMM), the Portland Digit Recognition Test (PDRT), and standard neuropsychological tests (Trail Making Test Parts A and B, and the Digit Symbol Coding subtest and Incidental Learning task of the WAIS-III). Three primary goals were: (a) compare the ability of the TOMM and the PDRT to correctly classify control participants and simulated malingerers, (b) examine the effect of different levels of information about brain damage (i.e., coaching) provided to simulators, and (c) assess the face validity of the measures. Participants were 96 college students randomly assigned to the following groups: controls, simulate brain damage without any additional information (i.e., no coaching), simulate brain damage with minimal coaching, or simulate brain damage with detailed coaching. After completing the tests, participants’ beliefs regarding the purpose of each test were obtained before and after they were informed that some tests were designed to detect feigned brain damage. Results indicated that the PDRT classified significantly more malingerers than the TOMM when using chance level responding; however, neither test identified malingerers at a satisfactory level. The TOMM classified significantly more malingerers than the PDRT when using published cutoff scores. The classification rate for the TOMM using cutoff scores was satisfactory (71% malingerers, 100% controls). In general, the control group performed better than the three malingering groups; the malingering groups did not differ from one another based on the levels of coaching. Both the PDRT and the TOMM have good face validity as memory tests. Once given a clue that some tests were designed to detect feigned brain damage, well over half of the participants identified all the tests as malingering measures. Potential explanations for the findings and directions for future research are discussed.
Acknowledgments

I would like to thank Stuart Hall, Ph.D. for serving as my mentor and committee chair. Dr. Hall’s knowledge, guidance, and incredible patience throughout the development of this thesis are greatly appreciated. In addition to teaching me valuable research skills, he has taught me a great deal about the clinical aspects involved in neuropsychology.

I would also like to thank the other members who served on my committee: John Klocek, Ph.D., Wendy Shields, Ph.D., and Annie Sondag, Ph.D. Moreover, I would like to thank John Caruso, Ph.D. for his assistance with statistical analyses. Even though Dr. Caruso was not serving on my committee, he never hesitated to help.

Additionally, I would like to acknowledge the following undergraduate research assistants for their dedication and hard work: Jamal Carney, Jared Chamberlain, Tyler Cherry, Nicolyn Fleming, Sara Laney, Cheyenne Laue, Yvonne Matoya, Jim Nevala, Elizabeth Porter, Lee Quentel, David Reardon, Grant Rodeway, Isaiah Sarault, Tom Seastone, and Erika Wimmer. Special thanks are extended to Erika Wimmer for her incredible leadership qualities and to Cheyenne Laue for always being willing to go beyond the call of duty.

Finally, I would like to thank my friends and family for their invaluable encouragement and support. They are very instrumental in my life as I pursue my goals.
# Table of Contents

Abstract .................................................................................................................................... ii  
Acknowledgments .................................................................................................................. iii  
Table of Contents ................................................................................................................... iv  
List of Tables ........................................................................................................................... viii  
List of Figures ......................................................................................................................... ix  
Introduction ............................................................................................................................. 1  
  Malingering in Neuropsychology ................................................................................ 2  
  Techniques Used to Detect Malingering ....................................................................... 3  
    Clinical Judgment ....................................................................................................... 3  
    Performance on Neuropsychological Tests ............................................................ 5  
    Patterns Across Tests .................................................................................................. 9  
    Specific Malingering Tests ....................................................................................... 11  
    Symptom Validity Techniques ................................................................................... 13  
  Coaching ........................................................................................................................... 17  
  Comparison of Malingering Tests ............................................................................... 20  
  Test of Memory Malingering ......................................................................................... 21  
  Purpose of the Present Study .......................................................................................... 25  
  Hypotheses ........................................................................................................................ 26  
Method ..................................................................................................................................... 29  
  Participants ......................................................................................................................... 29  
    Controls ........................................................................................................................... 30  
    Uncoached Malingersers ............................................................................................... 30
N. UCM Participants With Scores Above and Below Cutoff on the Brief Symptom Inventory (BSI) .......................................................... 100

O. Correlated Proportions Analyses for Comparing Classification Rates for the TOMM and the PDRT ......................................................... 101
List of Tables

1. College Student Norms and Percentage of Participants Elevated on the Brief Symptom Inventory (BSI) ................................................................. 42
2. Demographic Information for All Participants ............................................................. 43
3. Mean Number of Correct Responses on the TOMM .................................................... 44
4. Mean Number of Correct Responses on the PDRT ..................................................... 45
5. Mean Number of Seconds Participants Required to Complete the Trails Making Test Parts A and B ................................................................. 49
6. Mean Number of Correct Responses on the Digit Symbol Coding and Incidental Learning Pairing and Free Recall Tasks ................................. 50
7. Mean Effort and Success Ratings to Follow Instructions ........................................... 53
8. Correlation Coefficients Between Familiarity With Head Injury (HI) and Performance on the TOMM and the PDRT .................................. 55
9. Percentage of Participants Endorsing the Various Categories for Each Test on the PEQ1 ................................................................. 57
10. Percentage of Participants Identifying a Test as a Malingering Test .......................... 59
11. Certainty Ratings for Each Test Identified as a Malingering Measure ........................ 60
List of Figures

1. Percent of Participants Correctly Classified by the TOMM and PDRT Based on Chance Level Performance ................................................................. 47

2. Percent of Participants Correctly Classified by the TOMM and PDRT Based on Established Cutoff Score Performance .................................................. 48


4. Percentage of Participants Believing They Were Successful at Hiding Their Group Membership .......................................................... 54
Malingering is defined in the *Diagnostic and Statistical Manual of Mental Disorders-IV* (1994) as “the intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives” (V65.2, p. 683). Other researchers feel that this definition is too restrictive and view malingering as existing on a continuum. For example, Lipman (1962) identified the following four types of malingering: (a) invention, (b) perseveration, (c) exaggeration, and (d) transference. Invention malingering refers to cases in which people falsely claim the existence of symptoms that they have never experienced. Perseveration malingering occurs when individuals continue to claim the presence of symptoms that have since subsided. The third type of malingering, exaggeration, exists when the reported symptoms are worse than the genuine symptoms. This type of malingering is at times a result of the person fearing that their problem will go unnoticed if the symptoms are not exaggerated. Lipman’s final type of malingering is seen in transference malingerers when they falsely attribute real symptoms to an injury other than what actually caused the symptoms. Rogers (1997) further defined malingering on a continuum of mild, moderate, and severe depending on the degree to which the malingerers exaggerate or fabricate their deficits. Regardless of the type or extent of malingering, people vary in the degree to which they are consciously aware of their actions. That is, the motivation to malinger may be driven by unconscious or fully conscious factors (Travin & Protter, 1984).
Malingering in Neuropsychology

In the area of neuropsychological assessment, the issue of malingering has been shown to be very important. Researchers have demonstrated that individuals with normal neurological function can purposefully alter their scores on neuropsychological measures, appearing similar to patients with brain injuries. Results such as these have stimulated a large amount of research designed to develop techniques to detect malingerers in neuropsychological assessments.

Research aimed at detecting malingering in neuropsychology dates back to when Rey (1941) developed one of the first measures specifically designed to detect malingerers. The aim was for these malingering measures to be easy for patients while appearing difficult to malingerers. This testing paradigm allows differences to emerge between people responding honestly and malingerers. In addition to malingering tests, researchers have also determined that scores on standard neuropsychological measures can be useful in detecting malingering. While taking these tests, malingerers typically overestimate the severity of neurological deficits and produce patterns of scores that are distinguishable from the scores obtained from actual neurological patients (Benton & Spreen, 1961; Goebel, 1983; Heaton, Smith, Lehman, & Vogt, 1978).

Unfortunately, researchers using malingering tests and standard neuropsychological measures have shown that many of these techniques have notable limitations. For example, the tests often fail to distinguish between patients performing honestly (poor specificity) and people exaggerating deficits (poor sensitivity; Guilmette, Hart, Giuliano, & Leininger, 1994; Schretlen, Brandt, Krafft, & Van Gorp, 1991). Additionally, the purpose and ease of some of the tests are apparent to many individuals.
Techniques Used to Detect Malingering

Clinical Judgment

Malingers often exhibit behaviors that differ significantly from real patients. Malingering may be suspected when individuals show a disparity between neuropsychological performance and other aspects of their presentation. That is, an individual may be capable of functioning normally at his/her job but obtain scores on neuropsychological tests indicative of severe impairment (Cercy, Schretlen, & Brandt, 1997). Disparity between symptom complaints, objective test findings, and severity of injury is also common (Ruff, Wylie, & Tennant, 1993). Specifically, a person may present with symptoms that are indicative of a severe head injury while the medical test findings and description of the origin of injury suggest a mild injury. Furthermore, reports provided by the patients and their friends and family many times do not add up to a believable, consistent picture (Iverson, 1995).

In addition to the above indicators, malingerers may exhibit the following characteristics. They may report sudden onset of symptoms and exaggerate the symptoms to the extent of producing a bizarre portrayal of the condition. Malingerers may also show signs of frustration and aggravation during the evaluation (Iverson, 1995).
Additionally, they may act evasive, unfriendly, suspicious, and uncooperative. While responding to test questions, they may miss very easy items while correctly answering difficult ones, give up easily, and claim they do not know many of the answers. Finally, many patients may resist treatment, refuse employment, and blame all their problems on the disorder in question (Ruff et al., 1993).

Many of these indicators and characteristics rely largely on subjective clinical judgment. When an individual portrays certain characteristics, clinicians must determine if the individual's portrayal is representative of a neurological condition, malingering, or various extraneous factors that can influence how someone behaves during the interview and testing sessions. This judgment can at times be beneficial during the diagnostic process. For example, a clinician's opinion regarding the above indicators, provides additional data that can be integrated with objective test data. Specifically, a patient's behavior during the interview and information in their records can aid clinicians when they are attempting to understand various discrepancies and atypical test patterns indicative of malingering (Auerbach, 1992; Ruff et al., 1993). Moreover, some researchers have suggested that clinicians who are experienced in neurological disorders may be able to use their expertise to determine what symptoms are likely real and whether the medical history, test findings, and behavioral indicators provide a consistent picture (Auerbach, 1992; Pankratz & Binder, 1997).

Contrary to providing beneficial information, clinical judgment may be detrimental while the clinician is determining a patient's diagnosis. If a clinician believes that a person is malingering, then this belief can negatively affect rapport and bias the clinician's diagnostic decisions (Ruff et al., 1993; Zielinski, 1994). When clinicians are
biased, they may fail to consider alternative explanations for the individual’s behavior. Many things other than malingering, such as fatigue, depression, pain, anxiety, medication side-effects, and the neurological condition in question can influence how a person behaves during the interview and testing process (Ruff et al., 1993; Zielinski, 1994). As previously mentioned, experience may aid some clinicians while they are making diagnostic decisions. When clinicians do not have the necessary training, they may be unable to adequately distinguish between real symptoms and malingering behaviors (Brandt, 1988), leading to misdiagnoses that seriously affect patients’ lives (Zielinski, 1994). Finally, clinicians who depend on specific indicators to aid in the diagnostic process may fail to detect those malingerers who do not show such indicators. As a result of these issues, the data that are gathered based on clinical judgment should never be considered solely as signatures of malingering (Ruff et al., 1993; Zielinski, 1994).

Performance on Neuropsychological Tests

Researchers have demonstrated that simulators are capable of faking neurological deficits on some of the traditional neuropsychological tests, producing levels of impairment that are similar to brain injury patients. When considering the numerical score obtained by simulators and brain injury patients on neuropsychological tests, some researchers have had a difficult time distinguishing between the groups of participants on certain measures. In light of such findings, researchers have also investigated the type of responses that people provide on tests as a way to classify simulators and neurological patients. These researchers have found that although simulators are able to lower their scores to appear brain damaged, they often overexaggerate specific deficits and produce
patterns of scores on certain neuropsychological tests that differ from actual brain injury patients (Benton & Spreen, 1961; Goebel, 1983; Heaton et al., 1978; Suhr & Gunstad, 2000). The simulators' patterns often do not make neurological sense, reflecting the erroneous beliefs that many laypersons have about neurological conditions and their associated sequelae (Aubrey, Dobbs, & Rule, 1989). Throughout several decades, researchers have identified different patterns between simulators and neurological patients on tests measuring motor ability, attention, memory, sensory-perceptual functioning, language, IQ, visuospatial skills, and personality (Benton & Spreen, 1961; Bernard, 1990; Goebel, 1983; Hayward, Hall, Hunt, & Zubrick, 1987; Heaton et al., 1978; Mittenberg, Azrin, Millsaps, & Heilbronner, 1993).

Benton and Spreen (1961) conducted one of the initial studies showing that neuropsychological tests can be faked. They administered the Visual Retention Test to 47 college students, 23 medical patients without history of head injury or cerebral dysfunction, and 48 patients with cerebral damage. The researchers found that the college students and medical patients, serving as simulators, could be reliably distinguished from brain injury patients. In fact, the simulators tended to perform worse than patients with actual brain injuries. In addition, malingerers made more distortion errors while the brain injury patients made more omission and addition errors. This study showed that malingerers typically overestimate the severity of deficit (performing significantly worse than true brain injury patients) and present with atypical patterns of performance.

Heaton et al. (1978) conducted another landmark study and discovered that the level of functioning in simulators and head injury patients was similar on the Halstead-Reitan Battery and the Wechsler Adult Intelligence Scale (WAIS). However, while
simulators were able to lower their scores to be similar to head injury patients, several differences emerged between the test patterns produced by the two groups. Specifically, the malingerers obtained worse scores than the head injury patients on attentional, sensory-perceptual, and motor tasks. In contrast, the head injury patients did significantly worse on tasks measuring concept formation, general brain abilities, and tactile form discrimination and dexterity.

Heaton and colleagues (1978) also determined that the simulators had higher scores than the patients on the Minnesota Multiphasic Personality Inventory (MMPI). The simulators’ profiles showed elevations on the F scale and the following clinical scales: Hypochondriasis, Hysteia, Paranoia, Psychasthenia, Schizophrenia, and Social Introversion. Elevations on a malingerer’s profile appear to occur because of a tendency to overexaggerate negative symptoms. In contrast, neurological patients typically endorse symptoms that produce elevations that reflect bona-fide physiological and cognitive symptoms associated with their neurological conditions. Scales that are commonly elevated in neurological patients are Depression, Psychasthenia, and Schizophrenia (Cullum & Bigler, 1987).

Numerous additional studies have contributed to the findings that neuropsychological test performance differs between participants feigning neurological deficits and people responding honestly. For example, Goebel (1983) discovered that participants simulating various types of brain damage were unable to produce believable test patterns on tasks measuring a variety of cognitive abilities. Hayward et al. (1987) asked registered nurses to fake deficits suggestive of left fronto-temporal damage. Even though the nurses had years of experience working with neurological disorders, they
faked deficits that differed significantly from neurological patients, which resulted in atypical patterns of performance. Bernard (1990) discovered that participants simulating cognitive deficits on neuropsychological memory tests performed significantly lower than controls on 20 of 22 scores. Furthermore, Mittenberg et al. (1993) reported that their simulators scored one standard deviation below average on general memory tasks and two standard deviations below average on tasks measuring attention and concentration. This pattern of performance does not make neurological sense because it is necessary to first sufficiently attend to a task in order to remember the information for the memory task.

Researchers have examined serial position effects as another way to distinguish simulators and neurological patients. The serial position effect is observed when people recall items at the beginning (primacy effect) and the end (recency effect) of the list with higher frequency than items at the middle of the list, thus resulting in a “U” shaped learning curve (Rundus, 1971). Bernard (1991) found that simulators suppressed the primacy effect and closed head injury patients suppressed the recency effect on the Rey Auditory Verbal Learning Test. As expected the control participants demonstrated the typical “U” shaped curve. Contrary to the above study Bernard, Houston, and Natoli (1993) did not find significant serial position differences between the control and simulating participants. Finally, Wiggins and Brandt (1988) found that amnesiacs suppressed the primacy effect, but simulators produced the typical “U” shaped curve. These findings indicate that the serial position effect fails to provide a consistent way to distinguish between malingerers and people who respond honestly.
Researchers administering memory tests that utilize a recall and recognition format have demonstrated differences between simulators and head injury patients. Simulators consistently perform worse on recognition tasks than on recall tasks (Beetar & Williams, 1995; Bernard, 1990; Wiggins & Brandt, 1988). Often their recognition performance shows severe impairment while their recall performance is within normal limits. This pattern is directly opposite of what typically happens with normal individuals and people with neurological impairments. The simulators’ results are explained by two reasons. First, simulators do not realize that recognition performance is often intact in head injury patients; therefore, they overexaggerate their feigned deficits. Second, it is often difficult for simulators to determine what constitutes poor recall performance.

Patterns Across Tests

Patients often undergo multiple evaluations that include many of the same tests. Therefore, Cullum, Heaton, and Grant (1991) proposed that examining a person’s consistency of results across repeated evaluations is another useful technique for identifying feigned or exaggerated symptoms. Discrepancies occur when malingerers are not able to reproduce the same test patterns from one testing period to another. In order to reproduce the same patterns, malingerers must remember their previous responses and understand the sequelae of their feigned condition (Cullum et al., 1991; Owens, 1995). Additionally, malingerers’ performances often differ from what is typically observed when patients perform in an honest manner. Due to practice effects, retest scores obtained from real patients with stable neurological conditions (i.e., no new or developing neurological problems) tend to be the same or higher than previous test scores (Cullum et al., 1991).
The Halstead-Reitan Battery produces reliable, consistent scores when used in retest situations. Using this battery, Cullum et al. (1991) reported test results obtained from three litigating head injury patients. The patients did not have any evidence of complications that could cause deterioration in their functioning. All three patients produced variable test findings. Specifically, each patient’s scores increased on some measures and decreased on others across multiple testing periods. The variability in performance was determined to be inconsistent with what typically occurs in retest situations when patients with stable conditions perform to the best of their ability. Consequently, each patient’s performance was considered indicative of malingering.

Reitan and Wolfson (1997) administered the Halstead-Reitan Battery and either the original version of the WAIS or the Revised Edition (WAIS-R) to 40 head injury patients. Half of the patients were involved in litigation. Compared to their initial scores, nonlitigating patients obtained higher retest scores on each of the tests. Conversely, the litigating patients obtained lower retest scores on every test. The litigating patients’ retest scores did not show the improvement that typically occurs as a result of practice effects and previous exposure. This study and the research conducted by Cullum et al. (1991) show that responses that do not stay the same or do not improve across multiple testing periods are inconsistent with retest expectations and are possible signatures of malingering.

In summary, researchers have discovered that people can successfully fake deficits to the extent of appearing brain damaged; however, their test patterns and overall presentation often differ from real patients. Malingerers typically cannot fake patterns of deficit that make neurological sense. Nevertheless, clinicians’ ability to consistently
detect malingerers using clinical judgment and standard neuropsychological assessment batteries is poor (Faust, Hart, & Guilmette, 1988; Heaton et al., 1978). Consequently, neuropsychologists have been interested in developing specific tests for detecting malingered test performance.

**Specific Malingering Tests**

Researchers have developed measures specifically designed to detect malingering. The principle guiding the development of malingering tests was based on the notion that malingerers tend to overestimate deficits on tests, particularly when they believe the tests are difficult. Thus, researchers have attempted to design measures that are easy for patients with brain damage yet appear difficult. This test paradigm often allows differences to emerge between feigned and bona fide deficits. That is, malingerers tend to perform poorly, viewing the tests as difficult measures. In contrast, many patients are capable of obtaining high scores due to the ease of the tests (Iverson, Franzen, & McCracken, 1991). Examples of popular malingering measures that are commonly used are the Dot Counting Test (DCT) and the Fifteen-Item Memory Test (FIT).

The DCT, developed by Rey (1941) and described by Lezak (1995), is an instrument specifically designed to detect malingering behaviors. People are asked to count the dots as quickly as possible after the dots are presented on 3” X 5” cards. The total number of dots counted and response time are recorded. The first packet of six cards contains sets of ungrouped dots, and sets of grouped dots are presented on the second packet of six cards. Lezak (1995) suggested that malingering should be suspected when the time required to count the grouped dots exceeds the time needed to count the ungrouped dots. Furthermore, suspicions should surface when counting times are
inconsistent and not dependent on the number of dots presented. That is, when it takes people more time to count a lesser amount of dots, then malingering should be suspected (Binks et al., 1997). In order to test Lezak’s assertions several researchers have investigated the DCT as a way to detect malingering. Although some researchers were able to differentiate between malingerers and nonmalingerers (Binks et al., 1997), many other researchers have found the test to be inept at detecting malingerers (For a review see Greiffenstein et al., 1994; Hiscock et al., 1994; Rose, Hall, Szalda-Petree, & Bach, 1998).

The FIT was also developed by Rey (1964) and described by Lezak (1995). People are presented 15 items for 10 s and then they are asked to draw the items from memory. Using the FIT to assess patients, researchers have shown mixed results. Due to the similarities among the rows of stimuli, some researchers have found that the test is easy for many individuals (For a review see Beetar & Williams, 1995; Hiscock et al., 1994; Loring, 1995). Contrary to studies demonstrating the ease of the FIT, other researchers have discovered that patients with genuine neurological deficits do not consistently perform above the suggested cutoff scores, thus making them hard to distinguish from malingerers (For a review see Guilmette et al., 1994; Morgan, 1991; Rose et al., 1998; Schretlen et al., 1991).

In summary, these tests designed to detect malingering have been criticized for two reasons. First, malingerers sometimes detect the purpose of the tests and perform in the nonimpaired range. Second, these tests are challenging for some people with brain damage, making the distinction between brain damage patients and malingerers difficult.
Because both of these issues result in diagnostic limitations, researchers have focused their energy on developing more efficient measures.

Symptom Validity Techniques

Originally designed to detect deceptive sensory deficits, symptom validity techniques were adapted to be used in detecting feigned memory deficits (Binder & Pankratz, 1987; Hiscock et al., 1994; Pankratz, Fausti, & Peed, 1975; Ruff et al., 1993). In this paradigm, an initial stimulus, the target item, is presented to the individual. Then the target and a foil (two-alternative) are presented, and the person is required to choose one response (forced-choice) from the two possibilities. Because of this type of format, many symptom validity techniques are commonly referred to as two-alternative, forced-choice tests. Given that only two possible choices are provided, participants have a 50% chance of getting the correct answer when presented with a large number of trials (Binder & Willis, 1991; Hiscock & Hiscock, 1989). Accordingly, even if patients do not have any knowledge regarding the stimuli, which would be the case with total amnesia, they should get 50% of the questions correct simply due to the probability of chance.

Symptom validity techniques are used in two ways to detect malingers. First, performance is compared to chance levels. Malingers often overestimate deficits caused by a particular neurological disorder and believe that answering 50% of the questions correctly will prove they are not disabled. Therefore, they may perform below the level of chance, which is compelling evidence that they are intentionally responding incorrectly (Bernard et al., 1993; Loring, 1995). Unfortunately, researchers have found that not all malingers consistently perform this poorly (For a review see Binder, 1993; Binder & Willis, 1991; Hiscock et al., 1994; Iverson et al., 1991; Tombaugh, 1997). Consequently,
established cutoff scores are commonly used. These scores are above chance level of responding but lower than the scores produced by most brain injury patients and people without neurological disorders. Thus, malingerers are suspected when their test scores are lower than the scores obtained from patients (Hiscock et al., 1994; Tombaugh, 1997).

Binder and Pankratz (1987) suggested administering 100 forced-choice trials, making it more difficult for malingerers to monitor their performance. As they respond across trials, malingerers tend to reveal inconsistent response patterns. Binder and Pankratz used the symptom validity procedure with a 53-year-old woman and confirmed that the patient was performing significantly below chance when asked to specify whether she had been shown a black or yellow pen.

Hiscock and Hiscock (1989) believed that higher-functioning patients may react suspiciously to the simple task of identifying which pen had been presented. Therefore, they created a task that would be better suited for a broader range of patients. Instead of a pen as the target item, a 5-digit number is used in order to make the task appear more difficult. Hiscock and Hiscock's task consists of eight 5-digit numbers and 72 trials. Each target number is presented for 5 s, followed by response cards consisting of the target and a foil that are presented after a delay of either 5, 10, or 15 s. The task includes a trial-by-trial feedback component that allows malingerers to adjust their performance downward if they believe they are performing too well for a patient with bona fide deficits. Although the 5-digit numbers help create the impression that the task is difficult, in actuality people need to remember only the first or last digit in order to distinguish the target number from the foil item. Hiscock and Hiscock administered this test to a 45-year-old patient with
memory deficits and discovered that the patient obtained an overall score of 29% correct, a score significantly below chance level.

Binder (1990) further modified Hiscock and Hiscock's (1989) test by adding an interpolated task between stimulus presentation and recall. The Portland Digit Recognition Test (PDRT) is a 72 item, two-alternative, forced-choice test. The target items are 5-digit numbers that are presented at the rate of one digit per second. After this presentation, 5, 15, and 30 s delays occur during which time the participant counts backward from 20, 50, and 100 respectively. These delays are incorporated to make the test appear more difficult. Following the interpolated counting activity, the participant is presented a card containing the target item and a foil item (Binder, 1993; Binder & Willis, 1991; Iverson et al., 1994).

Litigating patients have been shown to perform significantly worse than nonlitigating patient on the PDRT. Additionally, simulators have been known to achieve the worst PDRT performance when compared to litigating and nonlitigating patients (Binder, 1993; Binder & Willis, 1991). Furthermore, Binder (1993) found that 26% of litigating patients fell below the established cutoff score while only 10% fell below chance level performance on the PDRT total score.

Rose, Hall, and Szalda-Petree (1995) developed a computerized version of the PDRT (PDRT-C) that has proven to be equivalent to the PDRT manual version when distinguishing between simulators, head injury patients, and controls. In addition to measuring participants' total number of correct responses, the PDRT-C assesses response latency in order to substantiate the notion that malingerers frequently take longer to answer questions than most people answering honestly (Iverson, 1995). An increase in
response latency may reflect malingerers' tendencies to overexaggerate their feigned deficits. Moreover, malingerers have to identify the correct answer and decide whether they want to respond accurately. If they decide to reject the correct answer, then they must take additional time to choose an incorrect response. Conversely, individuals responding honestly simply need to identify and choose the correct answer (Cercy et al., 1997).

Rose et al. (1995) confirmed that malingerers take longer to respond than controls on the PDRT-C. Furthermore, the researchers determined that classification errors were reduced by 32% when response latency and the PDRT-C total score were used to distinguish between head injury patients and malingering simulators. Moreover, classification rates for patients and simulators increased from 72% when using only the total score to 81% when both response latency and total score were used.

In summary, symptom validity techniques are one of the most sensitive tools for detecting malingering behavior (Guilmette et al., 1994). Malingerers can unequivocally be detected when their performance is lower than would be expected based on the probability of chance. Additionally, they may still be detected when their scores are above chance levels but below the scores obtained from neurological patients. Although these techniques have shown promise, they are typically boring for people to take and they often require a substantial amount of time to administer. These disadvantages limit their utility (Lezak, 1995). As a result, researchers continue to search for improved methods for detecting malingering behaviors.
Coaching

While studying malingering, many researchers have devoted time to understanding the impact that coaching has on test performance. Coaching involves providing information about a particular disorder (e.g., head injury) to simulators who are asked to feign deficits associated with the disorder in order to help the participants perform more similar to real patients (Rogers, Gillis, Bagby, & Monteiro, 1991). At times, some researchers assume that malingerers are detected because they lack knowledge about their expressed condition. When researchers rely on this assumption, malingerers who are knowledgeable about their condition may escape detection. Therefore, it is necessary to understand how this information affects test results (Martin, Gouvier, Todd, Bolter, & Niccolls, 1992).

The extent to which people are successful at faking may depend on the level of knowledge they possess about the manifestations of the disorder they are malingering. Individuals undergoing neuropsychological assessment can easily obtain knowledge about their condition. Libraries, support groups, physicians' and lawyers' offices, and the internet are among the many places people can go to get information. When people get information about their condition, they may produce test scores that are more similar to neurological patients than the test scores produced by people who do not obtain additional information. Thus, the similarities between the test scores of informed malingerers and patients make it more difficult for clinicians to distinguish between the two groups. Consequently, developing assessment strategies that can be used to identify malingerers who are knowledgeable about their condition should be an integral part of malingering research.
Although coaching techniques need to be further investigated, Ben-Porath (1994) cautioned researchers to consider the ethical issues involved with this type of research strategy. When researchers coach participants on specific test strategies and then publish this information, they risk undermining the assessment process because knowledge of these strategies can impact a psychological test’s integrity. Results from published coaching studies have the potential to help malingerers present more realistic test patterns, thus decreasing the probability of being detected. On the other hand, coaching studies provide clinicians and researchers beneficial information about how knowledge affects test results. Consequently, it is the researcher’s responsibility to weigh the costs and benefits and to determine the extent of information that is published.

Compared to simulators who are not provided with additional information, simulators who are given information about head injuries are more capable of producing test results that resemble real patients’ results. Kerr et al. (1990) provided simulators a magazine article regarding head injury effects, and they found that their participants’ performance closely resembled the performance of real patients. Martin et al. (1992) conducted a study and found that 38 of the 40 naïve (uncoached) simulators performed worse than the lowest brain injury patient’s score. The sophisticated simulators, who had received a warning to perform above chance levels and to miss more hard items than easy ones, scored higher than the naïve simulators. Although coaching improved the performance of the sophisticated malingerers, 60% still performed worse than the patients. Hiscock et al. (1994) administered seven measures to controls, coached simulators, and naïve (uncoached) simulators. They discovered that only the forced-choice Digit Memory Test distinguished between coached and naïve participants. These
studies indicate that additional information aided coached simulators during their presentation. However, coaching did not allow all individuals to escape detection.

Rose et al. (1995) used the PDRT-C to examine the effects of coaching. The coached malingerers received a description of head injury sequelae and were told to make their presentation “believable” and to not be “too obvious” by overly exaggerating impairments. The uncoached malingerers who were simply told to feign brain damage scored significantly below the coached malingerers, controls, and closed head injury patients. Using only the total correct score, the researchers correctly classified 83% of the uncoached malingerers and 47% of the coached malingerers. When the total correct score and response latency data were considered, classification rates decreased to 79% for the uncoached malingerers and increased to 70% for the coached malingerers. Overall, these results show that people who were provided information about head injuries were better able to escape detection but still many were detected.

The findings from a study conducted by Hall and Parker (1996) indicated that participants given very minimal coaching scored well above the cutoff score of 39 on the PDRT-C. The coaching manipulation consisted of simply telling people to make their presentation of brain damage “believable” without making it “too obvious”. No information was provided on how to accomplish this task, and they were not given information about the effects of head injuries. Interestingly, their scores were higher than those from a coached group in the previously mentioned study by Rose et al. (1995). The results of the Hall and Parker study are obviously disturbing because true malingerers will understand they at least need to make their presentation believable and not too obvious regardless if they invest time into obtaining information about the condition.
Johnson and Lesniak-Karpiak (1997) studied the effects of providing simulators a warning that there are ways to detect insufficient performance. Compared to simulators who were not warned, those receiving the warning obtained higher scores on verbal and general memory tasks. However, the two groups did not differ significantly on tasks measuring attention and concentration, motor ability, and visual memory. The results show that warning simulators that their malingered performance can be detected helps them improve their performance on some but not all measures.

In summary, results from studies employing coaching techniques show that people who are provided with information about the disorder in question are better able to portray a symptom picture similar to real patients than individuals who do not receive additional information. However, no coaching technique has allowed all simulating participants to remain undetected. Many malingerers are detected even after they have obtained information because they do not entirely understand how true patients respond to questions (Loring, 1995).

Comparison of Malingering Tests

Rose et al. (1998) compared the PDRT to three other malingering measures: The Rey Dot Counting Test (Lezak, 1995; Rey, 1941), the Nonverbal Forced Choice Test (Frederick & Foster, 1991), and the 21-Item Word List (Iverson et al., 1991). The four measures were administered to normal controls, head injury patients, and college students instructed to fake a head injury. Thirty college students were coached on the effects of head injuries, and 30 college students were uncoached. Considering all four tests, only 29% of the coached students were detected compared to 47% of the uncoached students. Only the PDRT correctly identified malingerers without misclassifying the head injury
patients. Moreover, the PDRT identified 47% of the coached malingerers compared to a detection rate of 33% or less of the coached individuals by each of the other three tests. When the coached and uncoached malingering groups were combined, the PDRT correctly classified 66% of the malingering participants and 89% of the head injury patients, resulting in a total hit rate of 74% compared to a total hit rate that was no better than chance by each of the other three tests.

Although Rose and colleagues (1998) demonstrated that the PDRT was superior at accurately classifying patients and in detecting both uncoached and coached participants, the test is not without disadvantages. Because of these disadvantages, some clinicians are hesitant to use the PDRT. First, the test takes approximately 45 min to 1 hr to administer. This lengthy administration time and the relatively easy nature of the PDRT cause many patients to become annoyed and bored with the entire assessment process, which may negatively affect their performance on many tests (Lezak, 1995). The lengthy administration time is also problematic given that neuropsychological evaluations are typically expensive and often take an entire day to complete. Thus, spending close to 1 hr to rule out malingering is not feasible. Finally, although most patients tend to find the PDRT easy, researchers have also demonstrated that the test is difficult for some patients with bona-fide deficits. Thus, patients and malingerers may obtain similar scores on the PDRT, making it difficult to distinguish between them (Binder & Willis, 1991).

Test of Memory Malingering

In response to the problems associated with current measures, researchers are searching for improved methods. The Test of Memory Malingering (TOMM) is a malingering instrument that shows promise. The test was designed to examine
recognition memory because literature has shown that people with neurological impairments generally obtain a high number of correct scores on picture recognition tasks. Thus, differences are likely to emerge between people with neurological impairments and malingerers who often overestimate impairments (Tombaugh, 1997).

The TOMM employs a two-alternative, forced-choice paradigm consisting of two learning trials and a delayed retention trial. On each learning trial, 50 drawings are presented at 3 s intervals; a two-choice panel containing a target and foil follows each trial. After finishing two trials and receiving feedback for each response, participants complete a retention trial following a 15 min delay period (Tombaugh, 1997).

Several studies have been conducted using the TOMM. An initial normative study was conducted on 405 volunteers ranging in age from 16 to 84 years. Participants were administered the original version of the TOMM and were asked to choose the target item when it was paired with three new drawings (foil items). The participants recognized 99% of the target items on Trial 2 and the retention trial, showing that people without neurological impairments are capable of achieving high scores on the TOMM. Moreover, people’s estimate regarding the number of pictures that they would be able to identify was significantly lower than their obtained scores, substantiating the notion that the TOMM has good face validity as a difficult test (Tombaugh, 1997).

Tombaugh (1997) formed an additional normative sample consisting of 70 volunteers ranging in age from 17 to 73 years. The new sample was administered the modified version of the TOMM, which is a forced-choice test that incorporates a feedback component following each response. Thus, the original four-alternative test was changed to the currently used two-alternative version. Analyses showed that the
participants’ scores for both normative samples were similar, indicating that the original and currently used versions are equivalent.

Researchers have conducted studies to determine how patients with a variety of diagnoses perform on the TOMM. Using cutoff scores and chance levels of responding, Rees, Tombaugh, Gansler, and Moczynski (1998) reported significant differences between hospital control patients, litigating traumatic brain injury (TBI) patients, and nonlitigating TBI patients. The researchers found that the litigating patients performed significantly worse than the nonlitigating TBI and hospital control patients who obtained nearly perfect scores on every trial.

Tombaugh (1997) further investigated patients’ performance by administering the TOMM to 161 nonlitigating patients with the following diagnoses: no cognitive impairment, cognitive impairment, aphasia, TBI, and dementia. Results indicated that on Trial 2 the patients with dementia averaged 92% correct. The other four groups of patients averaged scores that exceeded 97% correct, with the majority of the patients obtaining perfect scores. Consequently, patients with a variety of diagnoses are capable of achieving high scores on the TOMM.

Researchers have investigated how knowledge regarding head injuries influences people’s scores on the TOMM. Tombaugh (1997) tested 41 cognitively intact students. The 20 coached simulators were given 1 week to gather information on the effects of head injuries using any information available (e.g., books, lectures, friends, etc.). Then, the participants were asked to realistically simulate the effects of a head injury while completing several tests in addition to the TOMM. The simulators’ scores averaged only a few points higher than chance level, and none of the scores was higher than 42.
Average scores obtained from the 21 control students were nearly perfect. Thus, a cutoff score of 45 on Trial 2 correctly classified every individual. Moreover, Rees et al. (1998) tested 10 cognitively intact people and 18 TBI patients. Eight of the TBI patients were asked to exaggerate existing or previously experienced symptoms. They scored significantly lower than the other TBI patients and controls. Thus, personal knowledge and experience of TBI did not allow the patients’ exaggerated performance to go undetected.

On a limited basis, researchers have examined face validity for the TOMM. Tombaugh (1997) found that participants viewed the TOMM as a memory measure when it was administered along with the Hiscock test. Moreover, Rees et al. (1998) determined that the TOMM was viewed as a memory test when it was administered along with many other neuropsychological measures. These researchers examined participants’ beliefs regarding the face validity of the TOMM via a debriefing interview. Although the specifics regarding the exact questioning and techniques that the researchers used to obtain this information are not known, the results show that the TOMM has good face validity as a memory test and is not an obvious measure of malingering.

To summarize, based on the studies conducted thus far, the TOMM appears to offer improvements needed in malingering measures. The TOMM has a brief administration time, typically requiring only 15 min to complete. In addition, differences have been shown between litigating TBI patients and nonlitigating TBI patients. Moreover, neurologically impaired patients are capable of recognizing all 50 target pictures on the TOMM, distinguishing them from malingerers who often assume that they should perform poorly given the difficult appearance of the test. Thus, researchers
have achieved high classification rates between people performing honestly (high specificity) and people exaggerating deficits (high sensitivity). Researchers have also detected coached student simulators and patients who are knowledgeable about head injuries. Finally, people tend to view the TOMM as a difficult memory test; therefore, it has good face validity (Rees et al., 1998; Tombaugh, 1997). In conclusion, these many advantages make the TOMM a desirable instrument that warrants further research.

**Purpose of the Present Study**

There were three primary goals of this study. One goal was to compare the PDRT and the TOMM in their ability to distinguish between controls and three groups of simulated malingerers. When compared to three other malingering tests, the PDRT was superior to other popular malingering tests in classifying patients and simulators (Rose et al., 1998). To date few researchers have compared measures that are commonly used to detect insufficient performance. Therefore, it was useful to compare the PDRT and the TOMM.

A second goal was to manipulate the amount of coaching that was given to the simulated malingerers in order to determine how much information is necessary for people to realistically feign a head injury. Specifically, each malingerer received either no coaching, a warning to perform in a believable manner, or information about head injury effects. Thus, three types of instructions varying in the degree of coaching information were provided to the malingering participants. As previously reviewed, researchers have shown that coaching simulators on the effects of head injury helps those individuals perform more similar to head injury patients than when people do not receive coaching (Hall & Parker, 1996; Hiscock et al., 1994; Johnson & Lesniak-Karpiak, 1997; Kerr et
al., 1990; Martin et al., 1992; Rose et al., 1995, 1998). Therefore, it was necessary to evaluate levels of coaching when determining the efficacy of the PDRT and the TOMM. Currently, there are no known studies that have examined the effects of varying levels of coaching.

A third goal was to examine the face validity of the measures. As previously described, face validity for the TOMM has been investigated in a limited manner. The literature is lacking in information on the face validity for the PDRT and for the standard neuropsychological tests that were administered. Thus, this study explored in depth participants’ beliefs regarding the purpose of each test.

Hypotheses

1) The first hypothesis was based on participants’ expected test performance according to chance levels for the PDRT (36 correct) and all three trials on the TOMM (25 correct). Specifically, Uncoached Malingerers (UCM) were predicted to perform lower than chance levels on both tests. The Minimally Coached Malingerers (MCM), the Detailed Coached Malingerers (DCM), and the controls were expected to obtain scores that were higher than chance levels.

2) By considering chance level performance, it was hypothesized that the PDRT and the TOMM would be similar in their sensitivity to identify malingering participants while showing specificity in correctly classifying control participants.

3) The third hypothesis was based on participants’ expected performance according to established cutoff scores on the PDRT (39 correct) and the TOMM (45 correct). For the PDRT, it was hypothesized that the UCM would perform lower than the cutoff score. The MCM, DCM, and the controls were expected to perform higher than the
cutoff score. For all three trials on the TOMM, the three malingering groups were expected to perform lower than the cutoff score and controls were predicted to perform higher than the cutoff score.

4) When using the established cutoff scores, it was hypothesized that the TOMM would show better sensitivity and specificity, correctly distinguishing between malingering participants and control participants at a higher rate than the PDRT.

5) It was hypothesized that the degree of coaching would influence malingerers’ performance on the PDRT, the TOMM, and each of the standard tests. That is, the UCM were predicted to obtain fewer correct responses on all the tests than the MCM, DCM, and controls. The MCM were expected to respond to a significantly fewer number of correct items than the DCM who were expected to perform significantly lower (worse) than the control participants.

6) It was predicted that the control participants would obtain higher scores than the malingering participants, answering significantly more correct items on the PDRT, the TOMM, and all the standard tests.

7) It was further hypothesized that all groups of malingerers would perform worse on the Pairing portion of the Incidental Learning task than they would perform on the Free Recall portion.

8) A positive correlation was expected between familiarity with head injury sequelae and the scores obtained on the PDRT and the TOMM. That is, those participants who were more familiar with head injury sequelae were predicted to score higher on the two tests than those participants without this familiarity. It was hypothesized that this correlation would be highest for the UCM and the MCM.
9) Regarding the face validity for the measures, it was expected that the majority of participants would classify the TOMM as a memory measure. Face validity for the PDRT and the standard neuropsychological tests has not been studied. Therefore, no specific hypotheses were offered regarding these data.

10) When considering the mean certainty ratings regarding the degree to which participants were certain that a test was a malingering measure, it was hypothesized that the DCM would be the most certain about their decision regarding the true nature of the measures.
Method

Participants

Ninety-six college students were evenly divided and randomly assigned to one of four groups: (a) controls, (b) UCM, (c) MCM, and (d) DCM. The participants (40 men, 56 women, $M_{Age} = 19.58$) were recruited from an introductory psychology course at The University of Montana and received partial fulfillment of the experimental credit requirement for that course as compensation for participating. All participants were treated in accordance with the “Ethical Principles of Psychologists and Code of Conduct” (American Psychological Association, 1992).

Participants were screened for neurological and substance abuse problems as assessed by the medical history questionnaire (see Appendix A). Students who indicated a history of neurological and/or substance abuse problems were not included in the experiment.

Students were further screened for psychological problems using the Brief Symptom Inventory (BSI; see Appendix B). Criteria used to identify the degree of psychological distress in the students were obtained from a research study conducted with college students (Cochran, 1996). Given that college students tend to endorse many symptoms on the BSI (Cochran, 1985), the non-college student norms in the manual were deemed inappropriate for this study. Thus, the norms that Cochran (1996) collected on a college sample (33 men, 51 women, $M_{age} = 27.61, SD \ 3.72$) were used for the present study instead of the norms provided in the manual.

The following four groups of students participated in this experiment.
Controls. Twenty-four college students (Age 18-26, $M = 19.33$) were randomly assigned to the control group and were asked to perform to the best of their ability (see Appendix C). Rose et al. (1995, 1998) provided their control participants similar instructions.

Uncoached Malingerers (UCM). Another 24 college students (Age 18-28, $M = 19.54$) were randomly assigned to act as UCM, and they read the following scenario (see Appendix D). The scenario was modified from the scenarios used by Rose et al. (1995, 1998) and Tombaugh (1997).

You are about to take some cognitive tests that examine mental abilities such as attention, memory, thinking and reasoning skills, and your ability to think quickly. While responding to the tests, please pretend that you have experienced brain damage from a car accident involving a head-on collision. You hit your head against the windshield and were knocked out for 15 minutes. Afterwards, you felt “dazed” so you were hospitalized overnight for observation. Because the driver of the other car is at fault, you have decided to go to court to get money from the person responsible. During the next few months following the accident, the negative effects from your head injury disappear. Your lawsuit has not been settled yet, and your lawyer has told you that you may get more money if you look like you are still suffering from brain damage.

As you pretend to be this car accident victim, try to respond to each test as a patient who is trying to appear brain damaged in order to get money from the lawsuit. Thus, your performance on the tests should convince the examiner as well as the people involved in deciding the outcome of your lawsuit that you are still suffering from brain damage.

These students were not told how to accomplish their task and they did not receive information regarding the effects of head injuries. Thus, they represented the uncoached manipulation.

Minimally Coached Malingerers (MCM). A group of 24 college students (Age 18-32, $M = 19.88$) was randomly assigned to act as MCM. They received the same instructions and scenario as the UCM, but the following clue to make their portrayal of brain damage believable was added (see Appendix E).
In order to convince these individuals, your brain damage must be believable. These participants did not receive information about head injury sequelae; however, the clue to present a believable symptom picture represented a minimal level of coaching.

*Detailed Coached Malingerers (DCM).* Another 24 college students (Age 18-25, \( M = 19.58 \)) were randomly assigned to the DCM group. These students received the same instructions and scenario as the MCM, but they were also given the following information regarding head injury effects (see Appendix F). This information was adapted from the scenario that was used for the coached participants in the studies conducted by Rose et al. (1995, 1998).

Try to produce the most severe problems that you can without making it too obvious to the examiner. Major exaggerations, such as remembering absolutely nothing, are easy to detect. If the examiner does not believe that you have any problems you will not get any money for your head injury. People who have a head injury often have problems paying attention, cannot remember things as well, and do not learn things as easily as they did before their injury. They also think a little slower than they used to. Keep this in mind when taking the tests. Remember you are to try to mimic the performance of persons who are truly brain damaged.

This additional information described common head injury sequelae and emphasized that the participants' presentation must be believable and not obviously faked. This guidance represented the detailed-coached manipulation.

In light of the controversy surrounding coaching studies, the instructions used for the present study were limited to information that is widely available and easily accessed. Additionally, participants were not told to respond with at least a 50% accuracy rate. The PDRT and the TOMM employ a two-alternative, forced-choice paradigm making it possible to detect malingerers when their performance drops below the level of chance. Thus, informing participants that they should respond above chance levels would have invalidated the design of the two-alternative, forced-choice measures.
**Materials**

*Medical History Questionnaire.* The medical history questionnaire was used to assess participants' history of neurological and substance abuse conditions (see Appendix A). This questionnaire is a modified version of the questionnaire used by Rose et al. (1998).

*Brief Symptom Inventory (BSI).* The BSI is a 53-item self-report inventory that was used to assess psychological distress and symptom patterns in people (Administration Manual, Derogatis, 1993). Each item in this brief version of the SCL-90-R is rated on a five-point scale of distress, ranging from 0 (*not at all*) to 4 (*extremely*). Nine primary symptom dimensions and three global indices of psychological distress are assessed with the BSI. The inventory was originally normed on psychiatric and medical patients and community nonpatient members. As previously discussed, Cochran (1996) developed norms for college students, which were used for the present study.

*Portland Digit Recognition Test (PDRT).* The PDRT is a two-alternative, forced-choice test that was administered individually to each participant (Binder, 1990). Participants were verbally presented a 5-digit number at the rate of one digit per second. After the presentation, they were asked to count backwards from 20. This interpolated activity was interrupted after 5 s and then a 3×5 card was presented. The card had two 5-digit numbers, one being the target that was previously shown and one serving as the foil. The participants were asked to identify which item was shown and feedback was given as to whether they were right or wrong. This process occurred for 18 trials and the participants were praised if at least 12 of the 18 items were correctly identified. Then the examiner informed the participants that the task was going to get more difficult and they
were to count backwards from 50. This time the participants were not interrupted until a 15 s delay period had occurred. This process was repeated for another 18 trials and the same criterion was used to praise the participants. Following the 18 trials with the 15 s delay between each trial, participants were informed that the test was going to become even more difficult. The delay period was increased from 15 s to 30 s and the participants were asked to count backwards from 100 during the 30 s delay period. The test was concluded after another 36 trials had been completed. Items presented during the 5 s and 15 s delay trials were designated the “easy” items and the items presented during the 30 s delay were designated the “hard” items. Consequently, 36 “easy” and 36 “hard” items were presented prior to concluding the test. Even though items were designated “easy” and “hard”, the probability of getting the correct answer stayed constant for all items. That is, when participants were presented with one target and one foil, they had a 50% chance of correctly responding to each two-alternative item. The correct answer was located on top of the response card 50% of the time and on bottom the other 50% of the time.

Test of Memory Malingering (TOMM). The TOMM is also a two-alternative, forced-choice malingering test that was administered individually to each participant (Tombaugh, 1997). The test consisted of two learning trials and a delayed retention trial. Each learning trial included the same 50 line-drawn pictures of common objects. The drawings were presented at the rate of 1 drawing every 3 s and the order of picture presentation was different for the trials. A two-choice panel containing a target and a foil followed each learning trial and participants were asked to identify the drawing previously shown. A different foil was presented in a counterbalanced order for the trials.
A score of 0 to 50 was recorded following each trial and feedback was provided for every item. After a 15 min delay period that was filled with standard tests, retention was assessed.

*Trail Making Test (TMT) Parts A and B.* The TMT was administered during the 15 min delay period between Trial 2 and the retention trial for the TOMM (Reitan & Davison, 1974). Part A consisted of 25 circles printed on a piece of paper. A number from 1 to 25 was printed in the center of each circle. The participants were given instructions to connect the circles with a pencil as quickly as possible, beginning with the number 1 and working in a numerical sequence until they reached the number 25. Part B consisted of 25 circles that were numbered from 1 to 13 and lettered A to L. Participants were instructed to connect the circles as quickly as possible in sequence, alternating between numbers and letters (i.e., 1 to A, A to 2, 2 to B, B to 3, etc.). Performance was timed on both parts of the TMT. The examiner pointed out any mistakes that were made and the participants were asked to proceed from the point where the mistake occurred, correcting the error based on the feedback that was provided by the examiner. The scores for Part A and Part B were the number of seconds required to complete the tasks. The number of mistakes was not directly counted; however, mistakes increased the performance time, which affected the final score.

*Digit Symbol Coding Subtest.* The Digit Symbol Coding subtest from the Third Edition of the WAIS (WAIS-III) was also administered during the 15 min delay period for the TOMM (The Psychological Corporation, 1997). This subtest consisted of seven practice boxes and 133 test boxes. Each box contained a number in the upper portion and an empty space in the lower portion. On the same sheet of paper was a key that had nine
boxes numbered 1 through 9. Each number in the key was paired with a specific symbol. The examiner demonstrated the first few practice items and instructed the participants to complete the practice items by drawing the symbol that was paired with the number in the key. Once it was clear that the participants understood the task, they were instructed to fill in the test items without skipping any boxes while they worked. They were allowed 120 s to match as many numbers and symbols as possible. The score on this test was the total of correctly matched symbols and numbers.

*Digit Symbol Incidental Learning Task.* The Digit Symbol Coding subtest from the WAIS-III incorporates an additional task called the Digit Symbol Incidental Learning task (The Psychological Corporation, 1997). This task was administered during the 15 min delay period for the TOMM. The Pairing portion of the task consisted of two sets of nine boxes. Each set of boxes had a number in the upper portion of the boxes, ranging from 1 to 9. Participants were instructed to fill in the symbols that matched the given numbers. The symbol and number pairs were the same ones that the participants learned during the Digit Symbol Coding subtest; however, the key was not shown for the Incidental Learning task. Thus, this task was a measure of participants’ immediate ability to remember the appropriate matches. Additionally, the Free Recall portion of the Incidental Learning task was administered. Participants were instructed to write down all the symbols they remembered in any order they chose. The score for the Pairing portion of the task was the total of correctly matched numbers and symbols, and the score for the Free Recall portion was the total of correctly recalled symbols. Performance was not timed during any portion of the Incidental Learning task.
No known studies exist using the Incidental Learning task to examine test performance in malingerers. However, Hart, Kwentus, Wade, and Hamer (1987) administered the Incidental Learning task from the WAIS to healthy older adults, older adults experiencing depression, and older adults diagnosed with mild dementia. The researchers found that the performance of the people with depression was significantly lower than the healthy adults’ performance on the Pairing portion of the task, with no significant differences found between the two groups for the Free Recall task. Furthermore, the patients with dementia had significantly lower scores than the other two groups, obtaining few correct responses on both the Pairing and Free Recall portions of the task.

Post-Experimental Questionnaire 1 (PEQ1). All participants completed the PEQ1 (see Appendix G). Responses to the initial two items on the PEQ1 were used to determine the data that were included in the analyses. For the first item, participants were asked to paraphrase their instructions. Two raters determined whether participants’ written responses accurately reflected an understanding of their task. In order to establish the participants’ effort in following their instructions, a Likert-type item ranging from 1 (didn’t try at all) to 5 (tried very hard) was also included. Any data collected from participants indicating they correctly understood their instructions and tried at least moderately hard (i.e., a score of “3” on a 5-point Likert-type item) to follow the instructions were considered in the analyses.

Another Likert-type item, ranging from 1 (not at all successful) to 5 (very successful), was included to determine how successful the participants felt they were in accomplishing their task. Responses to a “Yes/No” question were used to assess whether
the participants believed they were successful in keeping the examiner from knowing what instructions they were following (i.e., control, UCM, MCM, and DCM). The participants were also asked to indicate their familiarity with head injury sequelae using a Likert-type item ranging from 1 (not at all familiar) to 5 (very familiar).

The last seven items on the PEQ1 were used to assess the face validity of the tests. Participants were asked to note their beliefs regarding the purpose of each test. Responses were categorized for the tests and were agreed upon by two independent raters (see Results section for the categories).

Post-Experimental Questionnaire 2 (PEQ2). An additional questionnaire was included to further assess the face validity of the seven tests (see Appendix H). This questionnaire contained instructions informing the participants that some of the tests were designed to measure feigned brain damage, while others are common tests used to assess cognitive abilities. Participants were asked to put a check by the tests that they believed were designed to evaluate feigned brain damage. Then they were asked to indicate the certainty of their beliefs on a Likert-type item ranging from 1 (not at all certain) to 5 (very certain). As previously reviewed, participants' initial beliefs regarding the purpose of each test were assessed using the PEQ1 described above. The PEQ2 was not provided to participants until after they sealed the PEQ1 in an envelope. Consequently, their initial responses were not contaminated by this additional information about the true intent of some of the tests.

Procedure

Prior to participating, students signed the informed consent form (see Appendix I). They were randomly assigned to groups according to the following procedure. One
envelope containing an instruction sheet and a second envelope containing the PEQ1 were prepared for each participant. An identification number was assigned to both envelopes and to all data for each individual. This number preserved the participants' anonymity and aided the examiners in organizing the data sheets. Each participant was tested individually.

The participants completed the BSI and the demographic sheet (age, completed years of education, and gender). Then they received a sealed envelope containing instructions representing one of the four different types of groups: (a) controls, (b) UCM, (c) MCM, and (d) DCM (see Appendixes C, D, E and F respectively). The instructions for the three malingering groups informed the participants that they would receive two additional credits if they successfully convinced the examiner that they had suffered a brain injury. In actuality, all participants received the additional credits in return for their participation. Given that students in the Psychology 100 courses are required to complete an experimental component, these additional credits provided an incentive for the malingering participants to follow their instructions. Incentives to malinger in the real world often include large monetary awards; therefore, it is important to provide some form of incentive in simulation studies.

A strict protocol was followed in order to keep the examiners blind to the individual conditions. For example, while the participants were reading the instructions, the examiner was not in the room. Each participant was instructed verbally as well as in writing to keep his or her group membership hidden at all times during the experiment. All participants were given 7 min to think about their instructions and were asked to seal
them in the envelope when they were finished reading them. Following the 7 min time period, the examiner returned to the examination room to proceed with the experiment.

The author or a research assistant administered the PDRT and the TOMM in counterbalanced order to each participant. The examiners were trained in test administration and followed a detailed protocol. During the 15 min delay period between Trial 2 and the retention trial for the TOMM, participants completed the TMT Parts A and B, followed by the Digit Symbol Coding subtest, and finally the Incidental Learning task. After completing all the measures, the participants received the PEQ1 in another sealed envelope (see Appendix G). While the participants were filling out this questionnaire, the examiner was not in the examination room. This provided further protection against the examiner discovering the participants' group membership as they paraphrased their instructions on this questionnaire. After completing the questionnaire, participants sealed it in the envelope and called the examiner into the room. The participants were then given the PEQ2 (see Appendix H), which was followed by a debriefing procedure (see Appendix J). Finally, participants were informed that they were being compensated for their efforts by receiving two additional experimental credits. The entire experimental session lasted an average of 2 hours.

Data Preparation and Analyses

A between-subjects' experiment was designed. The data analysis phase included 20 separate one-way anovas. In order to decrease the probability of making a Type I error, significant levels for anovas were corrected according to the Bonferroni using the following formula: .05 divided by the number of anovas. Thus, the individual alpha level used for all anovas was set to .001. The Tukey honestly significant difference comparison
test was used to further examine significant findings. The conventional .05 level of significance was utilized for post hoc analyses once significance had been initially achieved with the more stringent alpha level.

For the analyses in this experiment, there were 24 students in each of the four groups, for a total sample size of 96 participants. For this study, power exceeded .99 and the effect size was equal to 5.39. All data were analyzed using the Statistical Package for the Social Sciences (SPSS) version 10.0 (SPSS Inc., 1999).
Results

Five participants in the malingering groups did not indicate trying at least moderately hard (i.e., rating of “3”) to follow their instructions. Additionally, 2 participants in the malingering groups and 1 control participant did not indicate the required level of understanding regarding their instructions. These data were discarded and replaced by 8 additional participants whose effort was at the appropriate level and understanding of instructions was adequate.

Table 1 shows the norms obtained by Cochran (1996) as well as the percentage of participants in the present study who were elevated on the BSI. As previously reviewed, Cochran (1996) found that college students tend to have elevated scores on the BSI. Therefore, the 95th percentile from that study was used as the cutoff for the present study’s data. That is, students who fell at or above the 95th percentile were considered to have scores indicative of a significant degree of psychological distress. Using those criteria for the present study, 39 of 96 (41%) participants showed elevated scores on the BSI (i.e., elevations on at least 2 primary scales and/or an elevated Global Severity Index). Percentages of participants who were elevated on each of the scales are as follows: Somatization (17%), Interpersonal Sensitivity (27%), Obsessive-Compulsive (21%), Depression (11%), Anxiety (19%), Hostility (19%), Paranoid Ideation (22%), Phobic Anxiety (1%), Psychoticism (19%), and Global Severity Index (27%).

Statistical analyses for controls, DCM, MCM, and UCM whose scores were at/above or below the cutoff on the BSI are shown in Appendixes K, L, M, and N respectively. Independent samples t-tests were conducted to compare performance
Table 1

<table>
<thead>
<tr>
<th>BSI Scales</th>
<th>Normsa</th>
<th>Normsa</th>
<th>Presentb Study</th>
<th>Presentb Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M (SD)$</td>
<td>95th Percentile</td>
<td>% of Participants Elevated on 1 Scale</td>
<td>% of Participants Elevated on At Least 2 Scales and/or GSI</td>
</tr>
<tr>
<td>Somatization</td>
<td>.18 (.23)</td>
<td>.57</td>
<td>5</td>
<td>17</td>
</tr>
<tr>
<td>Interpersonal Sensitivity</td>
<td>.39 (.50)</td>
<td>1.00</td>
<td>3</td>
<td>27</td>
</tr>
<tr>
<td>Obsessive-Compulsive</td>
<td>.52 (.46)</td>
<td>1.33</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>Depression</td>
<td>.29 (.42)</td>
<td>1.17</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Anxiety</td>
<td>.33 (.38)</td>
<td>.83</td>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td>Hostility</td>
<td>.35 (.35)</td>
<td>1.00</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td>Paranoid Ideation</td>
<td>.29 (.44)</td>
<td>1.00</td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>Phobic Anxiety</td>
<td>.19 (.43)</td>
<td>1.2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Psychoticism</td>
<td>.22 (.42)</td>
<td>.80</td>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td>Global Severity Index (GSI)</td>
<td>.30 (.28)</td>
<td>.66</td>
<td>0</td>
<td>27</td>
</tr>
</tbody>
</table>

Note. a Means, standard deviations, and the 95th percentile normative data were gathered by Cochran (1996). b Percentages for current study include participants who were at or above the 95th percentile.
between participants with and without elevations on the BSI. No significant differences occurred on any of the measures when comparing participants whose BSI was above cutoff with participants whose BSI was below cutoff. Thus, participants were included in the analyses regardless of their score on the BSI.

Demographic Information

Table 2 contains the demographic information for the four groups of participants. Chi-square showed no significant gender differences, $\chi^2 (3, N = 96) = 3.09, p = .379$.

Group differences for age and education were analyzed by two separate anovas. No significant differences occurred for Age, $F(3, 92) = .21, p = .887$ or for years of education $F(3, 92) = .86, p = .465$.

Table 2

Demographic Information for All Participants

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Controls</th>
<th>DCM</th>
<th>MCM</th>
<th>UCM</th>
<th>$\chi^2$ or $F$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Men</td>
<td>11</td>
<td>13</td>
<td>8</td>
<td>8</td>
<td>3.09*</td>
</tr>
<tr>
<td></td>
<td>Women</td>
<td>13</td>
<td>11</td>
<td>16</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>$M$</td>
<td>19.33</td>
<td>19.58</td>
<td>19.88</td>
<td>19.54</td>
<td>(3, 92)+</td>
</tr>
<tr>
<td></td>
<td>$SD$</td>
<td>(1.79)</td>
<td>(1.79)</td>
<td>(3.40)</td>
<td>(2.11)</td>
<td>.21</td>
</tr>
<tr>
<td>Completed Years of Ed</td>
<td>$M$</td>
<td>12.38</td>
<td>12.67</td>
<td>12.38</td>
<td>12.29</td>
<td>(3, 92)+</td>
</tr>
<tr>
<td></td>
<td>$SD$</td>
<td>(.71)</td>
<td>(1.09)</td>
<td>( .92)</td>
<td>( .69)</td>
<td>.86</td>
</tr>
</tbody>
</table>

*Note. No significant gender differences exist. No significant differences occurred for age or years of education.*Chi-square. +F-Value.
**Malingering Tests**

*Test of Memory Malingering (TOMM).* The means and standard deviations for the number of correctly answered items on the TOMM for all three trials are presented in Table 3. Three separate one-way anovas identified significant group differences on Trial 1, \( F(3, 92) = 19.76, p = .000 \), Trial 2, \( F(3, 92) = 13.42, p = .000 \), and the Retention trial, \( F(3, 92) = 13.66, p = .000 \). Post hoc analysis using the Tukey honestly significant difference comparison indicated that the control group answered significantly more items correctly on all trials for the TOMM than the three malingering groups. The three malingering groups did not differ significantly from one another in their performance on any of the TOMM trials.

**Table 3**

*Mean Number of Correct Responses on the TOMM*

<table>
<thead>
<tr>
<th>Group</th>
<th>TOMM Trial 1</th>
<th>TOMM Trial 2</th>
<th>TOMM Retention Trial</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Controls</td>
<td>49.38(<em>a</em>)</td>
<td>.88</td>
<td>49.96(<em>a</em>)</td>
</tr>
<tr>
<td>DCM</td>
<td>33.25(<em>b</em>)</td>
<td>9.95</td>
<td>35.58(<em>b</em>)</td>
</tr>
<tr>
<td>MCM</td>
<td>30.12(<em>b</em>)</td>
<td>10.37</td>
<td>31.38(<em>b</em>)</td>
</tr>
<tr>
<td>UCM</td>
<td>34.13(<em>b</em>)</td>
<td>12.37</td>
<td>34.79(<em>b</em>)</td>
</tr>
</tbody>
</table>

*Note.* Means that do not share subscripts differ at \( p < .05 \) in the Tukey honestly significant difference comparison.
Portland Digit Recognition Test (PDRT): Total Score. Table 4 shows the means and standard deviations for the total number of correctly answered items on the PDRT. Anova identified significant group differences on the PDRT Total Score, $F(3, 92) = 39.20, p = .000$. Post hoc analysis using the Tukey honestly significant difference comparison showed that the control group answered significantly more items correctly than the three malingering groups. The three malingering groups did not differ significantly from one another on the PDRT.

Table 4

Mean Number of Correct Responses on the PDRT

<table>
<thead>
<tr>
<th>Test</th>
<th>PDRT Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>64.46&lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
<tr>
<td>DCM</td>
<td>41.63&lt;sub&gt;b&lt;/sub&gt;</td>
</tr>
<tr>
<td>MCM</td>
<td>37.04&lt;sub&gt;b&lt;/sub&gt;</td>
</tr>
<tr>
<td>UCM</td>
<td>40.13&lt;sub&gt;b&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

*Note.* Means that do not share subscripts differ at $p < .05$ in the Tukey honestly significant difference comparison.

Classification Rates Using the Malingering Tests

All three groups of malingering participants were combined into a single group of malingerers. The reason for creating one group of malingerers was that in clinical practice the goal is to be able to identify whether patients are malingering or whether they...
have bona fide deficits, regardless of whether the malingerers were coached for the
assessment process. In order to classify students as a malingering or a control participant,
the following rules were applied to the PDRT and the TOMM. Each participant’s
performance was compared to chance levels and to established cutoff scores for both
tests. For the PDRT Total Score, a score of 36 represents chance level performance and a
score of 39 represents cutoff score performance. For the TOMM, a score of 25 on Trial 2
represents chance level performance and a score of 45 on Trial 2 represents cutoff score
performance. Thus, participants who scored below chance levels or cutoff scores for the
PDRT and the TOMM were classified as malingering.

Figure 1 shows the classification rates for the PDRT and the TOMM according to
participants’ chance level performance. Correlated Proportions analysis was conducted to
determine the amount of participants who were correctly classified by only the PDRT or
the TOMM, by both tests, and by none of the two tests (see Appendix O). The analysis
showed that the PDRT correctly classified a significantly larger percentage of
malingering participants compared to the TOMM ($z = 2.31, p = .02$, two-tailed). On the
PDRT, 27 of 72 (38%) malingering participants were correctly classified. On the TOMM,
19 of 72 (26%) malingering participants were correctly classified. All control participants
were correctly classified by both tests (100% respectively). Overall, these findings show
that while the PDRT classified significantly more malingering participants, neither test is
effective at classifying participants when using chance level criterion (38% vs. 26%
respectively).
Figure 1

![Bar chart showing percent correctly classified by PDRT and TOMM for Malingerers and Controls.]

Percent of Participants Correctly Classified by the TOMM and PDRT Based on Chance Level Performance

Note. *p < .05.

Figure 2 shows that classification rates for the PDRT and the TOMM according to participants’ cutoff score performance. Again, Correlated Proportions analysis was conducted to determine the amount of participants who were correctly classified by only the PDRT or the TOMM, by both tests, and by none of the two tests (see Appendix O). The analysis showed that the TOMM correctly classified a significantly larger percentage of participants than the PDRT \( (z = 3.41, p = .001, \text{two-tailed}) \). On the PDRT, 35 of 72 (49%) malingering students were correctly classified. On the TOMM, 51 of 72 (71%) malingering participants were correctly classified. Both tests correctly identified all control participants according to cutoff score performance (100% respectively). These findings show that when using cutoff scores the TOMM classified significantly more
malingering participants than the PDRT, regardless of level of coaching (71% vs. 49% respectively).

Figure 2

Percent of Participants Correctly Classified by the TOMM and PDRT Based on Established Cutoff Score Performance

Note. *p < .05.

Standard Tests

Trail Making Test (TMT): Parts A and B. Table 5 shows the means and standard deviations for the total amount of seconds required for participants to complete each part of the TMT. Two one-way anovas indicated significant group differences on Part A, F(3, 92) = 8.79, p = .000 and on Part B, F(3, 92) = 8.30, p = .000. Post hoc analysis using the Tukey honestly significant difference comparison indicated that the control group showed significantly better performance than the malingering groups, completing both parts of
the TMT at a quicker speed. The three malingering groups did not differ significantly from one another in their performance on either part of the TMT.

Table 5

*Mean Number of Seconds Participants Required to Complete the Trails Making Test Parts A and B*

<table>
<thead>
<tr>
<th>Group</th>
<th>Trails A</th>
<th>Trails B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Controls</td>
<td>17.63&lt;sub&gt;a&lt;/sub&gt;</td>
<td>4.51</td>
</tr>
<tr>
<td>DCM</td>
<td>38.33&lt;sub&gt;b&lt;/sub&gt;</td>
<td>19.80</td>
</tr>
<tr>
<td>MCM</td>
<td>54.42&lt;sub&gt;b&lt;/sub&gt;</td>
<td>40.93</td>
</tr>
<tr>
<td>UCM</td>
<td>39.08&lt;sub&gt;b&lt;/sub&gt;</td>
<td>20.08</td>
</tr>
</tbody>
</table>

*Note.* Means in the same column that do not share subscripts differ at \( p < .05 \) in the Tukey honestly significant difference comparison.

*Digit Symbol Coding Subtest.* The means and standard deviations for the number of correctly answered items on the Digit Symbol Coding subtest are presented in Table 6. Anova revealed significant group differences on the subtest, \( F(3, 92) = 7.81, p = .000 \). Post hoc analysis using the Tukey honestly significant difference comparison showed that the control group answered significantly more items correctly when compared to all malingering groups. No significant differences occurred among the three malingering groups on the test.
Table 6

Mean Number of Correct Responses on the Digit Symbol Coding and Incidental Learning Pairing and Free Recall Tasks

<table>
<thead>
<tr>
<th>Group</th>
<th>Digit Symbol Coding</th>
<th>Incidental Learning Pairing</th>
<th>Incidental Learning Free Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>Controls</td>
<td>84.17&lt;sub&gt;a&lt;/sub&gt;</td>
<td>11.55</td>
<td>15.21&lt;sub&gt;a&lt;/sub&gt;</td>
</tr>
<tr>
<td>DCM</td>
<td>63.79&lt;sub&gt;b&lt;/sub&gt;</td>
<td>21.49</td>
<td>10.50&lt;sub&gt;b&lt;/sub&gt;</td>
</tr>
<tr>
<td>MCM</td>
<td>57.71&lt;sub&gt;b&lt;/sub&gt;</td>
<td>25.13</td>
<td>8.79&lt;sub&gt;b&lt;/sub&gt;</td>
</tr>
<tr>
<td>UCM</td>
<td>65.67&lt;sub&gt;b&lt;/sub&gt;</td>
<td>19.28</td>
<td>11.50&lt;sub&gt;b&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

Note. Means in the same column that do not share subscripts differ at $p < .05$ in the Tukey honestly significant difference comparison.

Digit Symbol Incidental Learning Task. The means and standard deviations for the number of correctly answered items for the Pairing portion and Free Recall portion of the test are also presented in Table 6. A one-way anova revealed significant group differences on the Pairing portion, $F(3, 92) = 8.35, p = .000$. Tukey honestly significant difference comparison indicated that the control group answered significantly more items correctly for the Pairing portion of the task than the three malingering groups. The three malingering groups did not differ significantly from one another in their performance on the Pairing portion of the Incidental Learning task. Anova showed significant group differences on the Free Recall portion, $F(3, 92) = 7.82, p = .000$. For the Free Recall portion, post hoc analysis using the Tukey procedure showed that the MCM group recalled significantly fewer items correctly than the UCM and the controls. Additionally,
the DCM recalled significantly fewer items correctly than the controls. There were no other significant differences among the four groups. The results for the Free Recall portion show that participants who received no coaching performed similar to control participants and people who received detailed coaching. Furthermore, participants without coaching performed significantly better than people who received minimal levels of coaching.

Figure 3 shows the percentage of correct responses that participants in each of the groups achieved on the Incidental Learning task Pairing and Free Recall portions. Paired samples t-tests were calculated for the four groups. These results indicated that participants’ performance on the Pairing portion was significantly lower than their Free Recall performance for the UCM, \( t(23) = -3.89, p = .001 \) (two-tailed), the MCM, \( t(23) = -4.96, p = .000 \) (two-tailed), the DCM, \( t(23) = -3.85, p = .001 \) (two-tailed), and controls, \( t(23) = -2.28, p = .032 \). These results show that regardless of whether participants are asked to respond honestly or to feign brain damage, they score worse on the Pairing task than they perform on the Free Recall task.

**Questionnaire Responses**

*First Post-Experimental Questionnaire (PEQ1).* The means and standard deviations for effort and success ratings are presented in Table 7. These ratings were analyzed by two separate one-way anovas. Significant group differences occurred for effort, \( F(3, 92) = 6.07, p = .001 \) and for success ratings, \( F(3, 92) = 35.40, p = .000 \). The Tukey honestly significant difference comparison revealed that the control participants rated their effort and their success at following their instructions as significantly better than the three malingering groups. These differences in ratings show that control
Figure 3

Performance on the Incidental Learning Task Pairing and Free Recall Portions

*Note.* Two-tailed paired samples t-tests showed significant differences between Pairing and Free Recall performance for each group, *p* < .05.

Participants felt the effort they put forth in following their instructions was greater than the effort reported by malingering participants. Additionally, these differences show that, compared to malingering participants, control participants felt more successful at accomplishing their experimental task. No significant differences occurred among the three malingering groups for effort or success ratings.
### Table 7

**Mean Effort and Success Ratings to Follow Instructions**

<table>
<thead>
<tr>
<th>Group</th>
<th>Effort Ratings</th>
<th></th>
<th>Success Ratings</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Controls</td>
<td>4.75$^a$</td>
<td>.60</td>
<td>4.75$^a$</td>
<td>.44</td>
</tr>
<tr>
<td>DCM</td>
<td>4.13$^b$</td>
<td>.74</td>
<td>2.83$^b$</td>
<td>.96</td>
</tr>
<tr>
<td>MCM</td>
<td>4.10$^b$</td>
<td>.75</td>
<td>3.10$^b$</td>
<td>.75</td>
</tr>
<tr>
<td>UCM</td>
<td>3.92$^b$</td>
<td>.78</td>
<td>2.63$^b$</td>
<td>.92</td>
</tr>
</tbody>
</table>

*Note.* Ratings were made on a 5-point scale according to participants' perceived effort (1 = didn't try at all, 5 = tried very hard) and success (1 = not at all successful, 5 = very successful). Means in the same column that do not share subscripts differ at $p < .05$ in the Tukey honestly significant difference comparison.

Responses to a "Yes/No" question on the PEQ1 were used to assess whether the participants believed they were successful in keeping the examiner from knowing what instructions they were following (i.e., controls, DCM, MCM, and UCM). Chi-square showed no significant differences among the four groups, $\chi^2(3, N = 96) = 4.02, p = .259$. Figure 4 shows that 96% of the control participants, 92% of the DCM, 79% of the MCM, and 92% of the UCM felt they were successful in keeping their group membership hidden.
Percent of Participants Believing They Were Successful at Hiding Their Group Membership

Note. Chi-square showed no significant differences among the four groups.

Pearson Correlation Coefficients, as shown in Table 8, indicated no significant correlation between familiarity with head injuries and participants’ scores on the PDRT Total Score. These results were found for the UCM \( r = -.19, p = .363 \), the MCM \( r = .19, p = .369 \), the DCM \( r = -.15, p = .484 \), and the controls \( r = -.02, p = .935 \).

Familiarity with head injuries and participants’ scores on Trial 2 for the TOMM were also not significantly correlated for the UCM \( r = -.20, p = .356 \), the MCM \( r = -.03, p = .899 \), the DCM \( r = .03, p = .900 \), and the controls \( r = .21, p = .316 \).
Table 8

*Correlation Coefficients Between Familiarity With Head Injury (HI) and Performance on the TOMM and the PDRT*

<table>
<thead>
<tr>
<th>Familiarity With HI</th>
<th>Controls (n = 24)</th>
<th>DCM (n = 24)</th>
<th>MCM (n = 24)</th>
<th>UCM (n = 24)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOMM Trial 2</td>
<td>.21 .316</td>
<td>.03 .900</td>
<td>-.03 .899</td>
<td>-.20 .356</td>
</tr>
<tr>
<td>PDRT Total Score</td>
<td>-.02 .935</td>
<td>-.15 .484</td>
<td>.19 .369</td>
<td>.19 .363</td>
</tr>
</tbody>
</table>

*Note.* No significant correlations occurred between familiarity with head injury and test performance on the TOMM and PDRT.
Seven items on the PEQ1 were examined to determine the face validity of the measures. All four groups of participants were combined for the analyses on this portion of the PEQ1 because the goal was to determine participants' beliefs regarding the purpose of each test, regardless of what instructions they received. That is, the primary question was whether the administered measures have good face validity as viewed by all people who take the tests.

The participants' responses were categorized for each test and agreed upon by two raters. The following categories were created: (a) Memory, (b) Higher Cognitive Skills, (c) Emotional Status, (d) Motor/Coordination, (e) Attention, (f) Memory and Attention, (g) Memory and Higher Cognitive, (h) Memory and Motor, (f) Don't Know/Don’t Remember, and (g) Other. Percentages are shown in Table 9 for the various categories endorsed for the tests. The majority of participants viewed the TOMM (94%), the PDRT (84%), and the Pairing portion of the Incidental Learning task (79%) as measuring some aspect of memory. Additionally, the TMT Part A (79%) and Part B (63%) and the Digit Symbol Coding subtest (51%) were most commonly viewed as tests used to measure higher cognitive skills (e.g., speed of information processing, association, thinking and reasoning, etc.). Interestingly, 48% of the participants believed the Digit Symbol Coding subtest was measuring some aspect of memory. Finally, the BSI was most commonly identified as a questionnaire that is administered to measure emotional status (75%).

These percentages represent participants' beliefs regarding the nature of the tests prior to receiving information that some of the tests may be malingering measures. The responses show that participants are generally accurate in identifying the nature of the standardized neuropsychological tests. Additionally, the TOMM and the PDRT have
good face validity as memory measures. Furthermore, with the exception of one individual who classified the BSI as a malingering measure (Other category), no response indicated that a test was designed to detect malingering.

Table 9

Percentage of Participants Endorsing the Various Categories for Each Test on the PEQ1

<table>
<thead>
<tr>
<th>Category</th>
<th>TOMM</th>
<th>PDRT</th>
<th>Trails A</th>
<th>Trails B</th>
<th>Dig Cod</th>
<th>IL Pair</th>
<th>BSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>92</td>
<td>69</td>
<td>11</td>
<td>11</td>
<td>42</td>
<td>73</td>
<td>5</td>
</tr>
<tr>
<td>Higher Cognitive</td>
<td>4</td>
<td>7</td>
<td>76</td>
<td>58</td>
<td>46</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td>Emotional Status</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>75</td>
</tr>
<tr>
<td>Motor</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Attention</td>
<td>2</td>
<td>8</td>
<td>2</td>
<td>19</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Memory/Attention</td>
<td>0</td>
<td>14</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Memory/Higher Cognitive</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Memory/Motor</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Don’t Know/Don’t Remember</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>

Note. n = 96 for each test except for Trails A in which n = 95.
Second Post-Experimental Questionnaire (PEQ2). As previously described, all four groups of participants were also combined for the analyses on the PEQ2. The primary issue was to determine which tests were viewed as malingering measures, regardless of participants’ instructions. Chi-square showed that the following tests were endorsed as malingering measures at significantly higher rates than would be expected by chance: the TOMM, \( \chi^2 (1, N = 96) = 40.04, p = .000 \), the TMT Part A, \( \chi^2 (1, N = 96) = 8.17, p = .004 \), the TMT Part B, \( \chi^2 (1, N = 96) = 9.38, p = .002 \), and the Digit Symbol Coding subtest, \( \chi^2 (1, N = 96) = 9.38, p = .002 \). Chi-square also indicated that the endorsement rates were not significant for the PDRT, \( \chi^2 (1, N = 96) = 2.67, p = .102 \) and the Incidental Learning Pairing portion, \( \chi^2 (1, N = 96) = 3.38, p = .066 \). The percentage of people who believed the BSI was not a malingering measure was statistically significant, \( \chi^2 (1, N = 96) = 12.04, p = .001 \). Table 10 shows the percentage of participants who classified each test as a malingering measure after receiving instructions that some tests were designed to detect malingering (see Appendix H for specific instructions). Overall, these findings indicate that participants viewed the TOMM (82%), TMT Parts A (65%) and B (66%), and Digit Symbol Coding (66%) as malingering measures. It appears the participants were just as likely to endorse the PDRT (58%) and the Pairing portion (59%) as malingering tests as they were to not endorse them. Finally, the majority of participants believed the BSI is definitely not a malingering test (68%). Thus, the TOMM was identified as a malingering test more than any of the other measures. However, three of the standardized tests were also viewed as malingering measures. In addition, all standard tests were viewed as malingering tests by over half of
the participants and a personality questionnaire was also thought to be a malingering test by approximately one third of the participants. Overall, these results indicate that college students are poor at identifying whether a measure was designed to identify malingering after receiving a clue that some tests were indeed malingering tests.

Table 10

**Percentage of Participants Identifying a Test as a Malingering Test**

<table>
<thead>
<tr>
<th>Test</th>
<th>% Yes</th>
<th>% No</th>
<th>$\chi^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOMM</td>
<td>82</td>
<td>18</td>
<td>40.04</td>
<td>.000</td>
</tr>
<tr>
<td>PDRT</td>
<td>58</td>
<td>42</td>
<td>2.67</td>
<td>.102</td>
</tr>
<tr>
<td>Trails A</td>
<td>65</td>
<td>35</td>
<td>8.17</td>
<td>.004</td>
</tr>
<tr>
<td>Trails B</td>
<td>66</td>
<td>34</td>
<td>9.38</td>
<td>.002</td>
</tr>
<tr>
<td>Dig Symbol Coding</td>
<td>66</td>
<td>34</td>
<td>9.38</td>
<td>.002</td>
</tr>
<tr>
<td>IL Pairing</td>
<td>59</td>
<td>41</td>
<td>3.38</td>
<td>.066</td>
</tr>
<tr>
<td>BSI</td>
<td>32</td>
<td>68</td>
<td>12.04</td>
<td>.001</td>
</tr>
</tbody>
</table>

*Note. $p < .05$.*

Table 11 indicates the mean certainty ratings and standard deviations for each test that participants identified as a malingering measure. For the following analyses, the four groups were not collapsed. Each group was examined separately in order to determine if the additional information provided to the DCM helped those participants feel more certain about their beliefs regarding the purpose of each test. Seven separate one-way anovas identified no significant group differences for mean certainty ratings on the
TOMM, $F(3, 75) = 1.24, p = .303$, PDRT, $F(3, 52) = .02, p = .995$, Trails A, $F(3, 58) = .79, p = .507$, Trails B, $F(3, 59) = .80, p = .499$, Digit Symbol Coding, $F(3, 59) = 1.73, p = .170$, Incidental Learning Pairing, $F(3, 53) = .97, p = .413$, and the BSI, $F(3, 27) = 2.25, p = .106$.

Table 11

*Certainty Ratings for Each Test Identified as a Malingering Measure*

<table>
<thead>
<tr>
<th>Test</th>
<th>Controls</th>
<th>DCM</th>
<th>MCM</th>
<th>UCM</th>
<th>$F$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOMM</td>
<td>3.61 (.120)</td>
<td>3.05 (.113)</td>
<td>3.55 (.94)</td>
<td>3.18 (1.05)</td>
<td>(3, 75)</td>
<td>.303</td>
</tr>
<tr>
<td>PDRT</td>
<td>3.31 (.95)</td>
<td>3.33 (1.44)</td>
<td>3.27 (1.03)</td>
<td>3.23 (1.01)</td>
<td>(3, 52)</td>
<td>.995</td>
</tr>
<tr>
<td>Trails A</td>
<td>3.43 (.94)</td>
<td>3.33 (.98)</td>
<td>3.13 (1.51)</td>
<td>2.90 (1.83)</td>
<td>(3, 58)</td>
<td>.507</td>
</tr>
<tr>
<td>Trails B</td>
<td>3.33 (1.14)</td>
<td>3.23 (.73)</td>
<td>2.82 (1.13)</td>
<td>3.07 (1.96)</td>
<td>(3, 59)</td>
<td>.499</td>
</tr>
<tr>
<td>Dig Symbol</td>
<td>3.36 (.84)</td>
<td>3.00 (1.00)</td>
<td>3.13 (1.45)</td>
<td>2.50 (1.04)</td>
<td>(3, 59)</td>
<td>.170</td>
</tr>
<tr>
<td>Coding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL Pairing</td>
<td>3.14 (.86)</td>
<td>2.50 (.85)</td>
<td>3.00 (1.17)</td>
<td>2.75 (1.00)</td>
<td>(3, 53)</td>
<td>.413</td>
</tr>
<tr>
<td>BSI</td>
<td>3.25 (.89)</td>
<td>2.60 (1.67)</td>
<td>1.75 (.89)</td>
<td>2.40 (1.26)</td>
<td>(3, 27)</td>
<td>.106</td>
</tr>
</tbody>
</table>

*Note.* Ratings were made on a 5-point scale according to participants' certainty regarding whether the test was a malingering measure (1 = not at all certain, 5 = very certain). Means for the four groups are not significantly different for any of the tests.
Discussion

One goal for the present study was to compare the PDRT and the TOMM in their ability to correctly distinguish between controls and three groups of simulated malingerers differentiated by the amount of coaching they received. These comparisons were based on participants’ chance level and cutoff score performance on the two tests.

Results showed partial support for the hypothesis regarding chance level performance among the four groups. All four groups of participants scored above chance level on both the PDRT and the TOMM. These results are consistent with many researchers’ findings: malingerers do not consistently perform below chance level on forced-choice tests (For a review see Binder, 1993; Binder & Willis, 1991; Hiscock et al., 1994; Iverson et al., 1991, Tombaugh, 1997). Contrary to the hypothesized results, the PDRT was more effective than the TOMM at correctly classifying malingering participants when chance level performance was considered, although both tests achieved very low classification rates for malingering participants. Both tests were equally effective at accurately classifying control participants.

Consistent with the hypothesis, the TOMM correctly classified a significantly larger percentage of malingering participants compared to the PDRT when the cutoff scores were used. Both tests correctly identified all control participants when using cutoff scores. Tombaugh (1997) demonstrated that using the established cutoff score on the TOMM correctly classified 91% of all patients, 99.9% of the cognitively intact volunteers, and 100% of the malingering participants (all received coaching). Rees et al. (1998) also found high classification rates for participants when using the cutoff score.
Interestingly, classification rates were lower for the malingerers in the present study when using the TOMM (71%).

As hypothesized, all three groups of malingerers scored below the established cutoff score and the control group scored above the cutoff score on the three TOMM trials. These data are consistent with Tombaugh (1997) and Rees et al. (1998) studies. The results showed partial support for the hypothesis that all groups, except the UCM, would perform higher than the cutoff score on the PDRT. The MCM group was the only group whose mean performance was lower than the established cutoff score for the PDRT. However, for all three malingering groups, test performance was within one to two points of the cutoff score. Rose et al. (1998) reported that their uncoached malingerers were several points below the cutoff score on the PDRT whereas their coached malingerers were within one point of the cutoff score. The head injury patients and control participants scored well above the cutoff score on the PDRT in the Rose et al. (1998) study.

Compared to the findings from the current study, Bianchini, Mathias, Greve, Houston, and Crouch (2001) demonstrated higher classification rates when using the PDRT and cutoff score performance. These researchers classified two groups of TBI patients: one group of probable malingerers (i.e., patients in which malingering was very likely) and one group of people who were unlikely to be malingering (i.e., patients who were not seeking compensation). All of the non-malingerers and 77% of the probable malingerers were classified by comparing the patients’ performance on the PDRT to the established cutoff score. These differences in classification rates between the present study and Bianchini et al.’s (2001) study appear to be due to the fact that many of the
simulated malingerers in this study performed better than expected, achieving scores above the cutoff score on the PDRT. Thus, a number of the malingerers escaped detection.

Overall, the findings support previous research showing that using chance level performance is not as effective as using cutoff score performance when classifying malingering participants (Bianchini et al., 2001; Rees et al., 1998; Tombaugh, 1997). Less than 40% of the malingering participants in the present study were identified when using chance level performance on both tests. Another issue previously reviewed is the fact that people often do not score lower than chance level on these tests (Binder, 1993; Binder & Willis, 1991; Hiscock et al., 1994; Iverson et al., 1991, Tombaugh, 1997). Thus, if researchers rely on chance level performance as a method for detecting malingerers, many individuals will escape detection.

Considering these results and previous research, the primary advantages that the TOMM has over the PDRT are brief administration time, better ability to detect malingerers using cutoff scores, and the fact that patients with severe deficits tend to do well on the TOMM (distinguishing them from malingerers). As previously reviewed, the PDRT has several notable shortcomings. First, dedicating approximately 1 hr simply to rule out malingering is not feasible given that neuropsychological evaluations are time consuming and expensive. Moreover, for many patients, the PDRT is relatively easy, which contributes to their boredom and annoyance with the assessment process (Lezak, 1995). At times the PDRT is difficult for some patients with bona-fide deficits, increasing the possibility that they will be inaccurately identified as malingerers (Binder & Willis, 1991). Thus, it seems that administering the TOMM, considering test patterns on
standard tests, and using clinical judgment would be beneficial techniques for detecting malingering signatures.

An additional goal was to examine in a thorough and systematic manner the impact that different levels of coaching have on test performance. Although malingering has been widely studied for decades, no studies have been published that examine coaching to the extent accomplished in the present study. Many researchers have demonstrated that coaching participants on the effects of head injury helps those participants perform better than when they do not receive coaching (Hall & Parker, 1996; Hiscock et al., 1994; Johnson & Lesniak-Karpiak, 1997; Kerr et al., 1990; Martin et al., 1992; Rose et al., 1995, 1998). Coaching has been shown to help participants have a clearer perception about how to portray a brain injury patient in a more believable manner. Thus, to extend existing research, the instructions in this study were systematically manipulated in such a way that provided the UCM with no coaching, MCM with one coaching clue (i.e., make performance believable), and DCM with head injury information and cautionary statements (i.e., to make performance believable and to not overexaggerate feigned brain damage).

In general, the results from this study do not support the hypotheses or previous research regarding coaching effects. The Free Recall portion of the Incidental Learning task was the only test in which differences emerged among the malingering groups. Further, it was the only test that controls did not do significantly better than all three malingering groups. Interestingly, on this test participants who did not receive any coaching performed similar to controls and to people who received information about head injuries, and they performed significantly better than those participants who
received minimal coaching. There was an insignificant trend for this pattern to occur among the malingering groups on all other tests as well. However, it is unclear why the participants without coaching performed similar to control participants on the Free Recall task. The following information includes potential explanations regarding the findings among the malingering groups.

Throughout this study, the coaching manipulation did not have the effect that was expected among the three groups of malingerers. That is, participants who received detailed information about head injury sequelae (DCM) and those who received the clue to make a believable presentation (MCM) did not perform better than those participants who received no additional information (UCM). In fact, there was a trend for the MCM to perform worse than the other three groups on the various measures.

It is important to consider potential reasons as to why the coaching manipulation did not result in significant differences between the malingerers. It appears that the subtle changes made to the scenarios used for the present study may have led to unexpected performance among the simulated malingerers. The scenarios for this study were created by combining portions of the scenarios used in the Tombaugh (1997) and Rose et al. (1998) studies. In certain aspects, the scenario used in Tombaugh (1997) study may potentially be perceived as indicating a less severe head injury than the scenario presented to the participants in the Rose et al. (1998) study. Additional differences, some more subtle than others, exist among the two scenarios. These results suggest that slight variations in the content of the scenarios can result in notable changes in participants’ performance. Thus, the content of the scenarios appears to be a critical feature of malingering research, which has not been previously studied.
The changes in scenarios likely led to the unexpected results that occurred among the UCM and the MCM. Overall, there was a trend for the UCM to perform better than was hypothesized. The UCM scored the same as the coached participants and 9 points higher than the uncoached participants on the PDRT total score in the Rose et al. (1998) study. On the three TOMM trials, the UCM performance was similar to the coached participants in the Rees et al. (1998) study. As previously reviewed, participants in the Rees et al. (1998) study were allowed 1 week to prepare their head injury presentation using any resources available to them. This preparation period served as a coaching manipulation for those participants. Again, the differences in the scenarios likely contributed to the inconsistent findings between these research studies.

The participants in the MCM group tended to perform worse than hypothesized on the various measures. The additional instruction to make their performance believable may have had the opposite effect as hypothesized. This clue was added to their instructions as a way to serve as a type of coaching. Participants in a study conducted by Hall and Parker (1996) were provided a clue to make their presentation believable without making it too obvious. Results showed that those participants performed better than the coached students in the Rose et al. (1998) study. Thus, for the present study it was hypothesized that the MCM would perform better than the UCM because they had been provided the clue to make their performance believable. Without any additional coaching information, this clue may have encouraged the MCM to overexaggerate their feigned symptoms. In other words, two possible interpretations may exist regarding how to make a believable presentation (a) to not overexaggerate deficits, leading to better test performance, and (b) to make sure the brain damage is seen by the examiner, leading to
worse performance. An overexaggeration of deficits in malingerers is consistent with previous research (Benton & Spreen, 1961; Goebel, 1983; Heaton et al., 1978). The large variations within the MCM group’s test performance seem to indicate that the participants were unclear about how to act.

The DCM were also told to make their performance believable. However, they were provided information about head injuries and warnings (i.e., do not be too obvious in the portrayal of brain damage and major exaggerations are noticeable). This additional information likely helped them in their portrayal of head injury compared to the MCM who did not receive this information. In general, the DCM performed similar to the coached participants in other research (Rees et al., 1998; Rose et al., 1998).

Another aspect that differs among malingering research studies involves the incentives provided to participants. For example, the present study offered two additional credits for participants if they successfully convinced the examiner that they had a brain injury. Rose et al. (1998) offered four additional credits. Rees et al. (1998) paid $50 to the one person in their study who presented a profile most typical of a brain injury patient. Tombaugh (1997) offered a specific amount of experimental credits to their participants; however, unlike some studies no additional credits were offered as an extra incentive. It is possible that these differences in incentives contribute to inconsistent findings among studies.

As hypothesized, all three groups of malingerers performed significantly worse on the Pairing portion of the Incidental Learning task than they performed on the Free Recall portion of the task. The control participants also performed significantly worse on the Pairing portion. In order to receive correct points on the Pairing portion people must
Malingering and Coaching

perform two steps: (a) recall the appropriate symbol and (b) correctly place the symbol with the associated number. The Free Recall portion allows more freedom in responses in that it requires the person to only recall the symbols without placing any restrictions on where to place the symbols. Thus, it is likely that the Pairing portion of the task is more difficult than the Free Recall portion as it was for all groups in the present study, contributing to the pattern found in test performance. It is interesting to note that the malingering groups showed a trend toward a greater disparity in Pairing vs. Free Recall scores than the control group. As previously mentioned, no other studies have examined malingers' performance on the Incidental Learning task. However, Hart et al. (1987) used the Incidental Learning task from the WAIS and found no significant differences between older adults with depression and older adults without depression for the Free Recall task. Furthermore, compared to people without depression, individuals with depression performed significantly worse on the Pairing portion of the task.

Contrary to the hypothesis, familiarity with head injury sequelae was not positively correlated with participants' performance on the PDRT or the TOMM. The participants in the present study varied in their familiarity with head injury, ranging from not at all familiar to very familiar (only a few people were very familiar with head injury sequelae). It appears that regardless of the amount of previous knowledge they had about head injuries, college students were unsure about how to portray a head injury patient. To examine the impact that preexisting knowledge has on malingers' performance, Rees et al. (1998) asked TBI patients to exaggerate existing deficits or deficits that had since subsided. These malingering patients obtained scores similar to college simulators,
showing that people with familiarity with head injury continue to overexaggerate deficits when asked to malinger on the TOMM.

Although the findings were not significant, the MCM group felt the least successful compared to the other three groups at hiding their membership from the examiner. These results may be due to the fact that the MCM group’s performance was overall worse than the other groups. That is, they altered their behavior more drastically than members in the other groups; therefore, they may have felt less successful at remaining inconspicuous (Bernard, 1991).

It needs to be noted that the control participants rated their effort and success at accomplishing their task as significantly better than the malingering participants with no significant differences occurring among the three malingering groups. In terms of effort, these findings are consistent with previous literature. Rose et al. (1998) found that their coached malingerers’ effort was similar to the uncoached malingerers’ effort and significantly worse than the effort reported by their controls and head injury patients. Bernard (1991) conducted a study using simulators and controls, producing results that were similar to the findings in the current study and to Rose et al.’s (1998) findings. Bernard (1991) hypothesized that the increased demand to perform in a specific manner with little preparation time led the simulators to rate their effort with lower scores. In the present study it appears that all malingering groups, regardless of the amount of coaching, may have viewed their task as difficult. Their perception of the task may have led them to view their effort and their success at accomplishing their task as worse than the controls.
The results of the present research have several implications. Although the coaching manipulation did not result in significant differences between malingering groups, the findings do shed light on the notion that variations in the content of the scenarios among studies lead to substantial differences in participants' performance. This is the first study to identify the critical role of the content of scenarios. The results show that researchers' conclusions that have been drawn for years about malingers' performance may be questionable and overall inconclusive due to differences in scenarios among the experiments.

Differences in the incentives provided to participants may also contribute to inconsistent results among experiments. Offering money as an incentive may influence participants' motivation on tests in a different manner than when experimental credits are offered. To further complicate matters, the amount of money and experimental credits is often very different among studies. Thus, such differences may further lead to inconsistent research findings.

Another goal was to examine the face validity of each administered measure. In spite of the fact that face validity has been considered to be a very important characteristic of a test, few researchers have explored this area (Bornstein, Rossner, Hill, & Stepanian, 1994). Researchers have shown that the TOMM has good face validity as a memory measure (Rees et al., 1998; Tombaugh, 1997). However, the literature is lacking regarding the face validity of other measures. Furthermore, no studies have explored face validity to the extent accomplished in the present experiment.

Chan, Schmitt, DeShon, Clause, and Delbridge (1997) showed that participants' test performance, their motivation, and their perceptions regarding the face validity of
measures all interact during testing procedures. Thus, depending on how people view
tests, their motivation may be affected as they take the various measures. These findings
emphasize that it is critical for the PDRT and the TOMM to have good face validity as
memory tests in order to enhance their efficacy as malingering measures. This also
suggests that it is important for standard neuropsychological tests to be viewed accurately
in regards to their true purpose.

The results of the present study showed that the PDRT and the TOMM have good
face validity as memory tests. Findings also showed that the Pairing portion of the
Incidental Learning task was most commonly viewed as measuring memory. The TMT
and the Digit Symbol Coding subtest were most commonly considered higher cognitive
tests. Furthermore, the BSI was most commonly identified as an emotional status
questionnaire. The results for the TOMM are consistent with the hypothesis as well as
Rees et al.’s study (1998) that showed the test has good face validity as a memory test
when it is embedded in a series of neuropsychological tests. Furthermore, these findings
are among the first to demonstrate that the PDRT has good face validity as a memory
measure and to show that standard neuropsychological tests are generally viewed
accurately.

Once given the clue that some of the tests may be malingering measures, over half
of the participants classified all the tests as malingering measures, with the TOMM being
classified by the largest percentage of people. Additionally, the BSI questionnaire was
classified as a malingering measure by over 30% of the participants. These findings
indicate that standardized tests and questionnaires may be viewed as malingering
measures. Furthermore, it appears that college students are poor at identifying whether a
measure was a malingering test after receiving a clue that some measures were designed to detect malingering.

It is important to consider why more students identified the TOMM vs. the PDRT as a malingering measure. One possibility is that the TOMM appears to be more similar to a memory test than the PDRT. Very few individuals classified the TOMM as measuring anything other than memory. Although the PDRT was initially classified by a large percentage of participants as a memory test, it was also classified as measuring higher cognitive skills and attention by a substantial percentage of people. It is likely that the interpolated activity (counting backwards) that is employed with the PDRT helped improve the face validity of this malingering test. Thus, the multi-purpose aspects of the PDRT may have led fewer individuals to choose it as a malingering measure.

The fact that memory impairment is a common complaint for real patients and malingerers (Cercy et al., 1997; Suhr & Gunstad, 2000) may be another potential explanation as to why the TOMM was endorsed more frequently than the PDRT as a malingering measure. That is, the participants may have believed that the test appearing most similar to a memory test was likely designed to be a malingering measure given that memory impairment is a common deficit following a head injury. As previously mentioned, the TOMM was endorsed as a memory test by 92% of the people whereas only 69% of the participants endorsed the PDRT as measuring only memory (an additional 14% also indicated that the PDRT was a measure of memory and attention). Researchers have shown that a large percentage of the general public are aware that memory impairment can occur following a concussion (Suhr & Gunstad, 2000). Additionally, Lees-Haley and Dunn (1994) demonstrated that more than half of their
introductory psychology students were capable of identifying several symptoms
associated with mild brain injury. These researchers’ findings may substantiate the
differences found in endorsement for the TOMM vs. the PDRT.

Participants’ certainty ratings regarding whether a test was a malingering measure
did not differ among the four groups of participants. Contrary to the hypothesized results,
the brief information that was provided to the DCM regarding head injury sequelae and
warnings about how to perform was not adequate for them to feel more confident in their
ability to identify the malingering measures. These findings may be due to the fact that in
general college students have limited knowledge regarding the purpose of assessment
measures, which likely contributed to their uncertainties.

This study was the first known attempt at examining face validity in this manner.
Interestingly, the participants were unable to discriminate between malingering and
standard tests on both face-validity questionnaires. Initially, the malingering tests were
viewed as memory tests. Conversely, after receiving a clue that some tests were designed
to detect feigned brain damage, the standard tests were viewed as malingering measures.
The differences between responses indicate that once people are made aware of the
existence of malingering measures, their views of tests change substantially yet remain
largely inaccurate. The participants were poor at identifying whether a test was a
malingering measure in spite of the fact that all tests were initially endorsed as measuring
memory, attention, speed of information processing, association, thinking, and/or
reasoning. These findings indicate that even if real patients are aware that malingering
tests exist, they are likely to be unsure about which tests were designed to detect
malingering.
In regards to these findings, it is crucial to gain a better understanding of peoples’ beliefs regarding the purpose of each test. It is important that malingering tests be viewed as standard tests. It is also important for standard tests to not be viewed as malingering measures. As previously mentioned, the manner in which people view tests impacts their test performance (Chan et al., 1997). This impact may not be limited to performance on individual tests but may affect the entire assessment process.

Finally, the results from the BSI showed that a significant amount of psychological symptoms does not negatively impact test performance. That is, participants who endorsed many psychological symptoms obtained similar scores on all the tests when compared to participants who did not express psychological distress, regardless of their group membership. These results indicate that the tests that were administered for the present study were not influenced by psychological symptoms in spite of the fact that the existence of psychological distress was substantial in this college population (41% of the participants had elevated scores indicative of psychological problems on the BSI). Elevations in college students may result from their perceptions regarding their participation in an experiment (i.e., they may be suspicious and/or nervous about the experimental process), stressful experiences related to being a college student (i.e., taking tests, meeting many deadlines, living away from home, etc.), and issues occurring in their personal lives.

These findings regarding the BSI have important clinical implications because many patients who present for neuropsychological assessments are also experiencing psychological problems in addition to their neurological sequelae. Thus, the findings may shed some light on how psychological distress impacts test results. In spite of these
findings, a caveat is in order. The degree and type of psychological distress that patients with neurological deficits and real malingerers experience may be different and in many cases more severe than the distress that our student participants endorsed. In fact, this appeared to be the case in the Hart et al. (1987) study. Hart et al. (1987) found that older adults experiencing depression performed worse than people without depression on the Pairing portion of the Incidental Learning task. Therefore, the impact of psychological sequelae must always be considered when interpreting test results.

Limitations of the Present Study

As with any simulation study, the generalizability of the results is limited. The incentive provided in this study is limited to experimental credits which is obviously not as appealing as the incentives in clinical cases (e.g., disability pay, workers’ compensation, veterans’ benefits, etc.). Furthermore, the participants were not allowed much time to prepare. On the other hand, patients who are feigning illnesses may learn to live their lives in ways that portray a true symptom picture. That is, they have an increased opportunity to practice living the role of a patient and are likely to learn how to present that role more realistically than simulators. An additional limitation lies in the fact that a clinical population was not included.

Future Directions

In the future, professionals in neuropsychology need to systematically examine the impact that subtle variations in instructions and incentives have on participants’ performance in order to gain a better perspective regarding which information tends to facilitate test performance and what seems to negatively influence performance. In order to provide consistent results that can be compared across future studies, it may be helpful
for researchers to create a standardized malingering scenario and incentive to be used across research laboratories.

The present study showed interesting trends regarding malingerers’ and control participants’ test performance on the Incidental Learning task, indicating that future researchers need to examine performance on this task as an additional method for detecting malingering. Additionally, more studies need to be conducted that include litigating TBI patients, nonlitigating TBI patients, and people with varying severities of TBIs. Furthermore, patients with different types of medical conditions and people with neurological impairments other than TBI need to be examined.

Assessing laypersons’ knowledge regarding head injury sequelae needs to be explored in future research, particularly as a way to shed light on face validity issues. It is unknown how many patients are familiar with head injury sequelae and the fact that some tests are designed as malingering measures. Essig, Mittenberg, Peterson, Strauman, and Cooper (2001) reported that most attorneys spend time educating their clients about neuropsychological assessment procedures. Topics that may be covered include malingering, content and names of commonly administered tests, information about head injury sequelae, testing strategies, and information that patients need to reveal during an interview. This literature substantiates the notion that the coaching and face validity aspects that were examined in this study need to be further incorporated into malingering research.
References


Hall, S., & Parker, L. R. (1996, April). *Test administration order and the Portland Digit Recognition Test-Computerized: Administer at will*. Poster session presented at the meeting of the Rocky Mountain Psychological Association, Park City, UT.


PLEASE FILL OUT THIS MEDICAL HISTORY FORM.

ALL INFORMATION YOU PROVIDE WILL BE HELD STRICTLY CONFIDENTIAL

1. Have you ever been diagnosed with any neurological condition?
   Yes  No
   If so, please list: _________________________________________

2. Have you ever had a blow to your head in which you were unconscious for longer than 30 minutes?
   Yes  No

3. Have you ever used drugs (e.g., alcohol, recreational drugs, Prescription medication) such that it caused you problems (e.g., difficulties at work, school, or home; other people believing that you have a problem with drugs; problems with the law; physical or psychological problems)?
   Yes  No
   If so, please list the specific problem(s) that your substance use caused:
   ____________________________________________________________
   ____________________________________________________________
## Appendix B

BSI

**INSTRUCTIONS:**
Below is a list of problems people sometimes have. Please read each one carefully, and circle the number to the right that best describes **HOW MUCH THAT PROBLEM HAS DISTRESSED OR BOTHERED YOU DURING THE PAST 7 DAYS INCLUDING TODAY.** Circle only one number for each problem and do not skip any items. If you change your mind, erase your first mark carefully. Read the example below before beginning, and if you have any questions please ask about them.

### Example

**HOW MUCH WERE YOU DISTRESSED BY:**

1. Bodyaches
   - **0=Not at all**
   - **1=A little bit**
   - **2=Moderately**
   - **3=Quite a bit**
   - **4=Extremely**

   **Example:**
   0 1 2 3 4

### Table

<table>
<thead>
<tr>
<th>Problem</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nervousness or shakiness inside</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faintness or dizziness</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The idea that someone else can control your thoughts</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling others are to blame for most of your troubles</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trouble remembering things</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling easily annoyed or irritated</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pains in heart or chest</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling afraid in open spaces or on the streets</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thoughts of ending your life</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling that most people cannot be trusted</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor appetite</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suddenly scared for no reason</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temper outbursts that you could not control</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling lonely even when you are with people</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling blocked in getting things done</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling lonely</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling blue</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling no interest in things</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix B

BSI

| HOW MUCH WERE YOU DISTRESSED BY: | 19. Feeling fearful | 20. Your feelings being easily hurt | 21. Feeling that people are unfriendly or dislike you | 22. Feeling inferior to others | 23. Nausea or upset stomach | 24. Feeling that you are watched or talked about by others | 25. Trouble falling asleep | 26. Having to check and double check what you do | 27. Difficulty making decisions | 28. Feeling afraid to travel on buses, subways, or trains | 29. Trouble getting your breath | 30. Hot or cold spells | 31. Having to avoid certain things, places, or activities because they frighten you | 32. Your mind going blank | 33. Numbness or tingling in parts of your body | 34. The idea that you should be punished for your sins | 35. Feeling hopeless about the future | 36. Trouble concentrating | 37. Feeling weak in parts of your body | 38. Feeling tense or keyed up | 39. Thoughts of death or dying | 40. Having urges to beat, injure, or harm someone | 41. Having urges to break or smash things | 42. Feeling very self-conscious with others | 43. Feeling uneasy in crowds, such as shopping or at a movie | 44. Never feeling close to another person | 45. Spells of terror or panic | 46. Getting into frequent arguments | 47. Feeling nervous when you are left alone | 48. Others not giving you proper credit for your achievements | 49. Feeling so restless you couldn’t sit still | 50. Feelings of worthlessness |
| 19. | 0 | 1 | 2 | 3 | 4 | 20. | 0 | 1 | 2 | 3 | 4 | 21. | 0 | 1 | 2 | 3 | 4 | 22. | 0 | 1 | 2 | 3 | 4 | 23. | 0 | 1 | 2 | 3 | 4 | 24. | 0 | 1 | 2 | 3 | 4 | 25. | 0 | 1 | 2 | 3 | 4 | 26. | 0 | 1 | 2 | 3 | 4 | 27. | 0 | 1 | 2 | 3 | 4 | 28. | 0 | 1 | 2 | 3 | 4 | 29. | 0 | 1 | 2 | 3 | 4 | 30. | 0 | 1 | 2 | 3 | 4 | 31. | 0 | 1 | 2 | 3 | 4 | 32. | 0 | 1 | 2 | 3 | 4 | 33. | 0 | 1 | 2 | 3 | 4 | 34. | 0 | 1 | 2 | 3 | 4 | 35. | 0 | 1 | 2 | 3 | 4 | 36. | 0 | 1 | 2 | 3 | 4 | 37. | 0 | 1 | 2 | 3 | 4 | 38. | 0 | 1 | 2 | 3 | 4 | 39. | 0 | 1 | 2 | 3 | 4 | 40. | 0 | 1 | 2 | 3 | 4 | 41. | 0 | 1 | 2 | 3 | 4 | 42. | 0 | 1 | 2 | 3 | 4 | 43. | 0 | 1 | 2 | 3 | 4 | 44. | 0 | 1 | 2 | 3 | 4 | 45. | 0 | 1 | 2 | 3 | 4 | 46. | 0 | 1 | 2 | 3 | 4 | 47. | 0 | 1 | 2 | 3 | 4 | 48. | 0 | 1 | 2 | 3 | 4 | 49. | 0 | 1 | 2 | 3 | 4 | 50. | 0 | 1 | 2 | 3 | 4 |
Appendix B

BSI

51. Feeling that people will take advantage of you if you let them.  51. 0 1 2 3 4
52. Feelings of guilt  52. 0 1 2 3 4
53. The idea that something is wrong with your mind  53. 0 1 2 3 4
Appendix C
Instructions for Controls

Please follow these instructions for the remainder of the experiment *without* letting the examiner know what you have been told to do! When you finish reading these instructions, sign at the bottom indicating that you have read them and understand your task. Then, place this signed sheet back into the envelope, seal it, and wait for the examiner to return. You will be asked about these instructions later on.

You are about to take some cognitive tests that examine mental abilities such as attention, memory, thinking and reasoning skills, and your ability to think quickly. Your task is to perform to the best of your ability, answering all questions in as honest a fashion as you can.

During the 7 minutes that the examiner is out of the room, remember to sign this sheet, place it in the envelope, and seal the envelope before the examiner returns.

Again, do not let the examiner know what these instructions have told you to do. Your performance on the tests should be the only way someone could figure it out.

Thank you very much for your participation.

I have read these instructions and will do my best to follow them for the remainder of the experiment.

(Signature)
Appendix D
Instructions for Uncoached Malingers

Please follow these instructions for the remainder of the experiment *without* letting the examiner know what you have been told to do! When you finish reading these instructions, sign at the bottom indicating that you have read them and understand your task. Then, place this signed sheet back into the envelope, seal it, and wait for the examiner to return. You will be asked about these instructions later on.

You are about to take some cognitive tests that examine mental abilities such as attention, memory, thinking and reasoning skills, and your ability to think quickly. While responding to the tests, please pretend that you have experienced brain damage from a car accident involving a head-on collision. You hit your head against the windshield and were knocked out for 15 minutes. Afterwards, you felt “dazed” so you were hospitalized overnight for observation. Because the driver of the other car is at fault, you have decided to go to court to get money from the person responsible. During the next few months following the accident, the negative effects from your head injury disappear. Your lawsuit has not been settled yet, and your lawyer has told you that you may get more money if you look like you are still suffering from brain damage.

As you pretend to be this car accident victim, try to respond to each test as a patient who is trying to appear brain damaged in order to get money from the lawsuit. Thus, your performance on the tests should convince the examiner as well as the people involved in deciding the outcome of your lawsuit that you are still suffering from brain damage. *If you succeed in convincing the examiner that you have suffered a brain injury, you will receive two additional experimental credits, for a total of 6 credits.*

During the 7 minutes that the examiner is out of the room, you may prepare for the examination. Remember to sign this sheet, place it in the envelope, and seal the envelope before the examiner returns. Again, do not let the examiner know what these instructions have told you to do. Your performance on the tests should be the only way someone could figure it out.

Thank you very much for your participation.

I have read these instructions and will do my best to follow them for the remainder of the experiment.

(Signature)
Appendix E

Instructions for Minimally Coached Malingerers

Please follow these instructions for the remainder of the experiment without letting the examiner know what you have been told to do! When you finish reading these instructions, sign at the bottom indicating that you have read them and understand your task. Then, place this signed sheet back into the envelope, seal it, and wait for the examiner to return. You will be asked about these instructions later on.

You are about to take some cognitive tests that examine mental abilities such as attention, memory, thinking and reasoning skills, and your ability to think quickly. While responding to the tests, please pretend that you have experienced brain damage from a car accident involving a head-on collision. You hit your head against the windshield and were knocked out for 15 minutes. Afterwards, you felt “dazed” so you were hospitalized overnight for observation. Because the driver of the other car is at fault, you have decided to go to court to get money from the person responsible. During the next few months following the accident, the negative effects from your head injury disappear. Your lawsuit has not been settled yet, and your lawyer has told you that you may get more money if you look like you are still suffering from brain damage.

As you pretend to be this car accident victim, try to respond to each test as a patient who is trying to appear brain damaged in order to get money from the lawsuit. Thus, your performance on the tests should convince the examiner as well as the people involved in deciding the outcome of your lawsuit that you are still suffering from brain damage. In order to convince these individuals, your brain damage must be believable. If you succeed in convincing the examiner that you have suffered a brain injury, you will receive two additional experimental credits, for a total of 6 credits.

During the 7 minutes that the examiner is out of the room, you may prepare for the examination. Remember to sign this sheet, place it in the envelope, and seal the envelope before the examiner returns. Again, do not let the examiner know what these instructions have told you to do. Your performance on the tests should be the only way someone could figure it out.

Thank you very much for your participation.

I have read these instructions and will do my best to follow them for the remainder of the experiment.

(Signature)
Appendix F

Instructions for Detailed Coached Malingerers

Please follow these instructions for the remainder of the experiment without letting the examiner know what you have been told to do! When you finish reading these instructions, sign at the bottom indicating that you have read them and understand your task. Then, place this signed sheet back into the envelope, seal it, and wait for the examiner to return. You will be asked about these instructions later on.

You are about to take some cognitive tests that examine mental abilities such as attention, memory, thinking and reasoning skills, and your ability to think quickly. While responding to the tests, please pretend that you have experienced brain damage from a car accident involving a head-on collision. You hit your head against the windshield and were knocked out for 15 minutes. Afterwards, you felt “dazed” so you were hospitalized overnight for observation. Because the driver of the other car is at fault, you have decided to go to court to get money from the person responsible. During the next few months following the accident, the negative effects from your head injury disappear. Your lawsuit has not been settled yet, and your lawyer has told you that you may get more money if you look like you are still suffering from brain damage.

As you pretend to be this car accident victim, try to respond to each test as a patient who is trying to appear brain damaged in order to get money from the lawsuit. Thus, your performance on the tests should convince the examiner as well as the people involved in deciding the outcome of your lawsuit that you are still suffering from brain damage. In order to convince these individuals, your brain damage must be believable. If you succeed in convincing the examiner that you have suffered a brain injury, you will receive two additional experimental credits, for a total of 6 credits.

Try to produce the most severe problems that you can without making it too obvious to the examiner. Major exaggerations, such as remembering absolutely nothing, are easy to detect. If the examiner does not believe that you have any problems you will not get any money for your head injury. People who have a head injury often have problems paying attention, cannot remember things as well, and do not learn things as easily as they did before their injury. They also think a little slower than they used to. Keep this in mind when taking the tests. Remember you are to try to mimic the performance of persons who are truly brain damaged.

During the 7 minutes that the examiner is out of the room, you may prepare for the examination. Remember to sign this sheet, place it in the envelope, and seal the envelope before the examiner returns. Again, do not let the examiner know what these instructions have told you to do. Your performance on the tests should be the only way someone could figure it out.

Thank you very much for your participation.

I have read these instructions and will do my best to follow them for the remainder of the experiment.

(Signature)
Appendix G

Post-Experimental Questionnaire 1

1. Please summarize the instructions you read at the beginning of this experiment.

2. Indicate how hard you tried to follow the instructions you were given at the beginning of the experiment by circling the number that best describes your effort.

   1  2 3 4 5
   Didn’t try at all Tried moderately hard Tried very hard

3. Indicate how successful you think you were in producing the results asked of you in the instructions by circling the number that best describes your success.

   1  2 3 4 5
   Not at all Successful Somewhat Successful Very Successful

4. Do you think you were successful in keeping the examiner from discovering your group membership (i.e., successful at concealing what your instructions told you to do)?

   Yes______   No______

5. Indicate how familiar you are with the effects that are often associated with a head injury by circling the number that best describes your familiarity.

   1  2 3 4 5
   Not at all Familiar Somewhat Familiar Very Familiar

6. What do you think the test with 50 different pictures was designed to measure? (Please write only one purpose for the test)
Appendix G

Post-Experimental Questionnaire 1

7. What do you think the test with the 5-digit numbers was designed to measure? (Please write only one purpose for the test)

8. What do you think the test with different numbers in circles (connected in dot-to-dot fashion) was designed to measure? (Please write only one purpose for the test)

9. What do you think the test with different numbers and letters in circles (connected in dot-to-dot fashion) was designed to measure? (Please write only one purpose for the test)

10. What do you think the test with different numbers and symbols (the test that provided a key matching symbols with numbers) was designed to measure? (Please write only one purpose for the test)

11. For the test with different numbers and symbols, what do you think the part with the numbers provided with blank squares was designed to measure? (Please write only one purpose for the test)

12. What do you think the questionnaire titled BSI containing 53 questions was designed to measure? (Please write only one purpose for the questionnaire)
Appendix H

Post-Experimental Questionnaire 2

It is possible that some of the tests you took today were designed to detect if someone is faking brain damage, while others are tests typically administered to test cognitive abilities such as memory, attention, and speed of information processing. Please put a check by any test that you took today that seemed as if it were designed to detect whether someone is faking brain damage. If you mark a test, please indicate how certain you are that the test was designed to detect faked brain damage by circling the number that best describes your certainty.

<table>
<thead>
<tr>
<th>Test Description</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 pictures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-digit numbers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numbers in circles (connected in dot-to-dot fashion)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Numbers and letters in circles (connected in dot-to-dot fashion)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matching numbers and symbols (the test that provided the number and symbol key)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matching numbers and symbols (the part that provided numbers with blank squares)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questionnaire titled BSI with 53 questions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix I

Informed Consent

Principal Investigator: Melody Huskey, Dept. of Psychology, The University of Montana, Missoula, MT 59812, 243-4521
Faculty Supervisor: Dr. Stuart Hall, Dept. of Psychology, The University of Montana, Missoula, MT 59812, 243-5667

Thank you for considering to participate in this study. This consent form may contain concepts that are unfamiliar to you. If the contents of this form are unclear, please ask the examiner to explain them to you. The purpose of this study is to investigate the effects of motivation on various tests. Additionally, we will be comparing the characteristics of the tests. If you agree to take part in this research study you will be administered some tests that examine mental abilities such as attention, memory, thinking and reasoning skills, and your ability to think quickly. Additionally, you will be asked to complete some questionnaires regarding problems that you may have had, some feelings you may have experienced, and drug and alcohol use. The session will last approximately two hours and will take place in the Skaggs building or the Clinical Psychology Center.

It is expected that the amount of discomfort you experience will be minimal. Some of the questions on the various questionnaires may cause you to feel uncomfortable or sad. Moreover, at times you may feel frustrated while completing the different tests. If these feelings occur, feel free to discuss them with the examiner and to contact the principal investigator or faculty supervisor at the numbers provided above.

Participating in this study will benefit you by providing you with experimental credits and giving you exposure to scientific research in psychology. Your participation will also provide beneficial information to professionals in the field of psychology.

The information you provide will be held strictly confidential by the research examiners (*see limits of confidentiality below). Your name will not be marked on the test answer sheets and questionnaires. However, if you agree to participate in this study, you will need to sign this form, which will be kept locked up and separate from all testing and questionnaire materials. We will have you note your age, gender, and years of education, but this personal identification information will not be attached to this form that contains your name. You will be assigned a participant number that will be used to help us keep your data sheets organized. The information that you provide will be read only by the principal investigator (Melody Huskey), the faculty supervisor (Dr. Stuart Hall), and the research assistants. Your test and questionnaire responses will be kept a minimum of 5 years after the study has ended; however, this sheet containing your name and phone number will be destroyed at the conclusion of the study. *There are conditions under which confidentiality may be breached. If you indicate wanting to harm yourself or someone else, then this informed consent form will be given to Dr. John Klocek (licensed clinical psychologist) who will contact you. Because of this, we also require that you provide your name and phone number below.

Name (print) ___________________________ Phone ___________________________
Appendix I

Informed Consent

Although there is minimal risk associated with your participation in this study, The University of Montana requires that the following paragraph be included in all consent forms:

“In the event that you are injured as a result of this research you should individually seek appropriate medical treatment. If the injury is caused by the negligence of the University or any of its employees, you may be entitled to reimbursement or compensation pursuant to the Comprehensive State Insurance Plan established by the Department of Administration under the authority of M.C.A., Title 2, Chapter 9. In the event of a claim for such injury, further information may be obtained from the University’s Claims representative or University Legal Counsel. (Reviewed by University Legal Counsel, July 6, 1993).”

Your participation in this study is entirely voluntary, and you may withdraw without penalty or any negative consequences. If you choose to withdraw, all your records will be destroyed, and the data you provided will not be used in this study.

If you have questions about this study now or during this session, please ask the examiner. Additionally, you may contact the principal investigator (Melody Huskey, 243-4521) if you continue to be unclear about the study. We will not be able to give you extensive feedback regarding your responses; however, you will be provided with additional information at the conclusion of the study. If you have any questions regarding your rights as a research subject, you may contact Dr. Rudbach (Chair of the IRB, 243-6670).

I have read the above description of this study and have been informed of the benefits and risks involved. All of my questions have been answered to my satisfaction, and I have been provided contact information for the principal investigator and the faculty supervisor in the event that I have concerns or questions in the future. By signing below I voluntarily agree to participate in this study and give my consent to the examiners to use the information I provide for the purposes of this experiment. I understand I will receive a copy of this consent form.

___________________________
Printed Name of Subject

___________________________
Subject’s Signature Date

___________________________
Examiner’s Signature Date
Appendix J

Debriefing Statement

Thank you for participating in this experiment. The purpose of this study was to examine the ability of different tests to detect if someone is faking a brain injury. Another goal of the study was to see if providing different amounts of information about brain injuries would affect how someone performs on these tests. Some subjects were simply asked to perform to the best of their abilities.

All subjects in this experiment played a very important role. The results from this experiment will enhance our knowledge about how people perform on these tests and whether or not they are giving their best effort. We thank you for your time in helping with our research.

We request that you please not discuss the details of this experiment with anyone who may be participants. Providing people details before completing the study may contaminate our results. However, by providing the information on this debriefing statement, the purpose will be thoroughly explained to all subjects at the conclusion of the experiment. Furthermore, participants will be allowed to ask questions following completion of the experiment.
### Appendix K

Control Participants With Scores Above and Below Cutoff on the Brief Symptom Inventory (BSI)

<table>
<thead>
<tr>
<th>Test</th>
<th>Controls (SD)</th>
<th>Controlsb (BSI)</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOMM Trial 1</td>
<td>49.26 (.93)</td>
<td>49.80 (.45)</td>
<td>-1.23</td>
<td>.230</td>
</tr>
<tr>
<td>TOMM Trial 2</td>
<td>49.95 (.23)</td>
<td>50.00 (.00)</td>
<td>-.51</td>
<td>.619</td>
</tr>
<tr>
<td>TOMM Retention</td>
<td>50.00 (.00)</td>
<td>50.00 (.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDRT Total</td>
<td>64.47 (4.21)</td>
<td>64.40 (3.29)</td>
<td>.04</td>
<td>.971</td>
</tr>
<tr>
<td>Trails A</td>
<td>16.95 (3.79)</td>
<td>20.20 (6.46)</td>
<td>-1.47</td>
<td>.155</td>
</tr>
<tr>
<td>Trails B</td>
<td>41.74 (9.67)</td>
<td>46.60 (13.94)</td>
<td>-.92</td>
<td>.370</td>
</tr>
<tr>
<td>Digit Symbol</td>
<td>83.42 (10.83)</td>
<td>87.00 (15.05)</td>
<td>-.61</td>
<td>.549</td>
</tr>
<tr>
<td>Coding</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incidental Learn</td>
<td>14.84 (3.48)</td>
<td>16.60 (.89)</td>
<td>-1.10</td>
<td>.283</td>
</tr>
<tr>
<td>Incidental Learn Free Recall</td>
<td>8.05 (1.13)</td>
<td>8.00 (.71)</td>
<td></td>
<td>.923</td>
</tr>
</tbody>
</table>

*Note.* Groups that are designated as "BSI" obtained scores at or above the cutoff. No t values are available for TOMM Retention Trial because the standard deviations for both groups are 0. a_n = 19. b_n = 5. Two-tailed independent t-tests showed no significant differences for the tests.
Appendix L

DCM Participants With Scores Above and Below Cutoff on the Brief Symptom Inventory (BSI)

<table>
<thead>
<tr>
<th>Test</th>
<th>DCM&lt;sup&gt;a&lt;/sup&gt; M (SD)</th>
<th>DCM&lt;sup&gt;b&lt;/sup&gt; M (SD)</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOMM Trial 1</td>
<td>33.47 (10.26)</td>
<td>32.89 (10.01)</td>
<td>.14</td>
<td>.894</td>
</tr>
<tr>
<td>TOMM Trial 2</td>
<td>35.00 (11.05)</td>
<td>36.56 (11.79)</td>
<td>-.33</td>
<td>.748</td>
</tr>
<tr>
<td>TOMM Retention</td>
<td>35.00 (10.73)</td>
<td>37.11 (13.81)</td>
<td>-.42</td>
<td>.679</td>
</tr>
<tr>
<td>PDRT Total</td>
<td>42.53 (10.91)</td>
<td>40.11 (12.11)</td>
<td>.51</td>
<td>.618</td>
</tr>
<tr>
<td>Trails A</td>
<td>36.40 (18.92)</td>
<td>41.56 (21.96)</td>
<td>-.61</td>
<td>.549</td>
</tr>
<tr>
<td>Trails B</td>
<td>69.13 (39.03)</td>
<td>73.89 (28.23)</td>
<td>-.32</td>
<td>.754</td>
</tr>
<tr>
<td>Digit Symbol Coding</td>
<td>64.27 (20.15)</td>
<td>63.00 (24.82)</td>
<td>.14</td>
<td>.892</td>
</tr>
<tr>
<td>Incidental Learning Pairing</td>
<td>10.93 (4.99)</td>
<td>9.78 (3.27)</td>
<td>.62</td>
<td>.544</td>
</tr>
<tr>
<td>Incidental Learning Free Recall</td>
<td>6.53 (1.36)</td>
<td>6.44 (1.42)</td>
<td>.15</td>
<td>.880</td>
</tr>
</tbody>
</table>

Note. Groups that are designated as “BSI” obtained scores at or above the cutoff. <sup>a</sup>n = 15. <sup>b</sup>n = 9. Two-tailed independent t-tests showed no significant differences for the tests.
Appendix M

MCM Participants With Scores Above and Below Cutoff on the Brief Symptom Inventory (BSI)

<table>
<thead>
<tr>
<th>Test</th>
<th>MCM&lt;sup&gt;a&lt;/sup&gt;</th>
<th>MCM&lt;sup&gt;b&lt;/sup&gt;</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOMM Trial 1</td>
<td>32.82 (10.66)</td>
<td>27.85 (9.96)</td>
<td>1.18</td>
<td>.250</td>
</tr>
<tr>
<td>TOMM Trial 2</td>
<td>33.91 (12.63)</td>
<td>29.23 (11.79)</td>
<td>.94</td>
<td>.359</td>
</tr>
<tr>
<td>TOMM Retention</td>
<td>31.64 (13.89)</td>
<td>27.77 (13.16)</td>
<td>.70</td>
<td>.492</td>
</tr>
<tr>
<td>PDRT Total</td>
<td>37.09 (10.12)</td>
<td>37.00 (11.57)</td>
<td>.02</td>
<td>.984</td>
</tr>
<tr>
<td>Trails A</td>
<td>51.18 (43.72)</td>
<td>57.15 (40.00)</td>
<td>-.35</td>
<td>.730</td>
</tr>
<tr>
<td>Trails B</td>
<td>91.55 (54.74)</td>
<td>94.08 (51.88)</td>
<td>-.12</td>
<td>.909</td>
</tr>
<tr>
<td>Digit Symbol</td>
<td>57.91 (27.93)</td>
<td>57.54 (23.67)</td>
<td>.04</td>
<td>.972</td>
</tr>
<tr>
<td>Coding</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incidental Learning Pairing</td>
<td>9.27 (5.82)</td>
<td>8.38 (4.74)</td>
<td>.41</td>
<td>.684</td>
</tr>
<tr>
<td>Incidental Learning Free Recall</td>
<td>5.64 (3.04)</td>
<td>6.08 (1.50)</td>
<td>-.46</td>
<td>.649</td>
</tr>
</tbody>
</table>

*Note.* Groups that are designated as “BSI” obtained scores at or above the cutoff. <sup>a</sup>n = 11. <sup>b</sup>n = 13. Two-tailed independent t-tests showed no significant differences for the tests.
Appendix N

UCM Participants With Scores Above and Below Cutoff on the Brief Symptom Inventory (BSI)

<table>
<thead>
<tr>
<th>Test</th>
<th>UCM&lt;sup&gt;a&lt;/sup&gt;</th>
<th>UCM&lt;sup&gt;b&lt;/sup&gt;</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOMM Trial 1</td>
<td>34.83 (13.97)</td>
<td>33.42 (11.11)</td>
<td>-.28</td>
<td>.786</td>
</tr>
<tr>
<td>TOMM Trial 2</td>
<td>33.83 (16.60)</td>
<td>35.75 (12.96)</td>
<td>-.32</td>
<td>.756</td>
</tr>
<tr>
<td>TOMM Retention</td>
<td>33.17 (15.05)</td>
<td>36.92 (15.00)</td>
<td>-.61</td>
<td>.547</td>
</tr>
<tr>
<td>PDRT Total</td>
<td>41.83 (12.16)</td>
<td>38.42 (11.07)</td>
<td>.72</td>
<td>.479</td>
</tr>
<tr>
<td>Trails A</td>
<td>38.00 (20.48)</td>
<td>40.17 (20.53)</td>
<td>-.26</td>
<td>.798</td>
</tr>
<tr>
<td>Trails B</td>
<td>63.33 (17.95)</td>
<td>85.92 (36.94)</td>
<td>-1.91</td>
<td>.070</td>
</tr>
<tr>
<td>Digit Symbol</td>
<td>67.58 (15.38)</td>
<td>63.75 (23.09)</td>
<td>.48</td>
<td>.637</td>
</tr>
<tr>
<td>Incidental Learning Pairing</td>
<td>12.58 (4.72)</td>
<td>10.42 (5.93)</td>
<td>.99</td>
<td>.333</td>
</tr>
<tr>
<td>Incidental Learning Free Recall</td>
<td>7.58 (1.16)</td>
<td>6.83 (1.90)</td>
<td>1.17</td>
<td>.256</td>
</tr>
</tbody>
</table>

*Note.* Groups that are designated as "BSI" obtained scores at or above the cutoff. <sup>a</sup>n = 12. <sup>b</sup>n = 12. Two-tailed independent t-tests showed no significant differences for the tests.
Appendix O

Correlated Proportions Analyses for Comparing Classification Rates for the TOMM and the PDRT

**Chance Level Rates**

<table>
<thead>
<tr>
<th>PDRT</th>
<th>TOMM</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10(^a)</td>
<td>17(^b)</td>
</tr>
<tr>
<td></td>
<td>43(^c)</td>
<td>2(^d)</td>
</tr>
<tr>
<td>Total</td>
<td>53</td>
<td>19</td>
</tr>
</tbody>
</table>

**Note.** \(^a\)Classified only by the PDRT. \(^b\)Classified by both the PDRT and the TOMM. \(^c\)Classified by none of the tests. \(^d\)Classified only by the TOMM.

\[ z = \frac{d - a}{\sqrt{a + d}}. \]

**Cutoff Score Rates**

<table>
<thead>
<tr>
<th>PDRT</th>
<th>TOMM</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3(^a)</td>
<td>32(^b)</td>
</tr>
<tr>
<td></td>
<td>18(^c)</td>
<td>19(^d)</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>51</td>
</tr>
</tbody>
</table>

**Note.** \(^a\)Classified only by the PDRT. \(^b\)Classified by both the PDRT and the TOMM. \(^c\)Classified by none of the tests. \(^d\)Classified only by the TOMM.

\[ z = \frac{d - a}{\sqrt{a + d}}. \]