University of Kentucky Doctoral Dissertations                    Graduate School

2007

# APPLYING REACTION TIME (RT) AND EVENT-RELATED POTENTIAL (ERPS) MEASURES TO DETECT MALINGERED NEUROCOGNITIVE DEFICIT

Victoria Louise Vagnini
*University of Kentucky*, nnagn0@uky.edu

ABSTRACT OF DISSERTATION

Victoria Louise Vagnini

The Graduate School

University of Kentucky

2007

APPLYING REACTION TIME (RT) AND EVENT-RELATED POTENTIAL (ERPS)
MEASURES TO DETECT MALINGERED NEUROCOGNITIVE DEFICIT

—————————————————————

ABSTRACT OF DISSERTATION

—————————————————————

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Arts and Sciences
at the University of Kentucky

By
Victoria Louise Vagnini

Lexington, Kentucky

Co-Directors: Dr. David T. R. Berry, Professor of Psychology

and Dr. Yang Jiang, Assistant Professor of Behavioral Science

Lexington, Kentucky

2007

ABSTRACT OF DISSERTATION

APPLYING REACTION TIME (RT) AND EVENT-RELATED POTENTIAL (ERPS)
MEASURES TO DETECT MALINGERED NEUROCOGNITIVE DEFICIT

This study examined the ability of reaction time (RT) and Event-Related
Potentials (ERP) to detect malingered neurocognitive deficit (MNCD)in two new tasks
compared to the TOMM ($N = 47$).  Honest (HON), malingering (MAL), and traumatic
brain injury (TBI) groups were compared on accuracy, RT and ERP measures.  Overall,
the Test of Memory Malingering (TOMM) accuracy was the most effective at classifying
groups (hit rate = 100%).  Several non-TOMM accuracy variables and RT variables
reached hit rates in the range of 71%-88%.  The TOMM RT variable had an unlimited
time for participants to respond and was the most successful RT variable compared to the
Old/New and Repetition Priming tasks that had a short time limit for participants to
respond (approximately 1.5 seconds).  The classic "old/new effect" RT pattern was
evident for both the HON and TBI groups with significantly faster RTs for old items
compared to new items.  A logistic regression was employed to see if a RT and/or ERP
variable added any unique prediction power in detecting malingering.  The frontal-
posterior ERP difference score had unique prediction power to detect malingering when
classifying MAL vs. TBI (hit rate = 86%).  In the Old/New task, ERP responses of HON
produced greater activity in the frontal region compared to the posterior region. The
opposite trend was found in TBI (posterior activity > frontal) and MAL showed no
significant difference.

KEYWORDS: Malingering, ERP, RT, Neurocognitive Deficit, TOMM-C

Victoria Louise Vagnini

July 16, 2007

APPLYING REACTION TIME (RT) AND EVENT-RELATED POTENTIAL (ERPS)
MEASURES TO DETECT MALINGERED NEUROCOGNITIVE DEFICIT

By

Victoria Louise Vagnini

David T. R. Berry
Co-Director of Dissertation

Yang Jiang
Co-Director of Dissertation

David T. R. Berry
Director of Graduate Studies

July 16, 2007

RULES FOR THE USE OF DISSERTATIONS

DISSERTATION

Victoria Louise Vagnini

The Graduate School

University of Kentucky

2007

APPLYING REACTION TIME (RT) AND EVENT-RELATED POTENTIAL (ERPS)
MEASURES TO DETECT MALINGERED NEUROCOGNITIVE DEFICIT

_____

DISSERTATION

_____

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Arts and Sciences
at the University of Kentucky

By
Victoria Louise Vagnini

Lexington, Kentucky

Co-Directors: Dr. David T. R. Berry, Professor of Psychology

and Dr. Yang Jiang, Assistant Professor of Behavioral Science

Lexington, Kentucky

2007

ACKNOWLEDGMENTS

I would like to thank my dissertation committee for their guidance and editorial comments. Thank you to David T. R. Berry and Yang Jiang, the directors of my dissertation, and to Brian Gold, Suzanne Segerstrom, and Anders Anderson.

I would also like to thank my mother, Nettie Vagnini, for her love and support without which I would not be where I am today.  Finally, I would also like to give thanks to God who I believe is behind all good things in this world.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter One
Introduction

*Background*

       With the increasing popularity of crime and espionage television shows like *Law & Order* or *Alias*, the use of lie detection tests such as the polygraph is now common knowledge. An interrogator may threaten a suspect with a polygraph test as a way to "prove" the subject's guilt. Often, the threat of having to take this test is enough to have the alleged criminal confess to the crime. While television is rarely an accurate reflection of real life, the rise in law enforcement TV shows has increased public awareness of tests designed to discern truth from lies in a way that appears to be more accurate than determining guilt or innocence based on a "gut feeling".

       The polygraph is a lie detection test that has been used with increasing regularity in situations like employment screening or to ensure the integrity of people who know top secret information. A polygraph device detects changes in heart rate, breathing, etc. during responses. Increases in these autonomic responses are thought to reflect anxiety. It logically follows that if a person is asked a question that would indicate his/her guilt (e.g., "Did you steal from Bank X?"); autonomic changes from measurements taken during baseline questions (e.g., "Is your name X?") should take place. The autonomic fluctuations occur due to increased anxiety, even if the outward appearance of the person is unchanged. By measuring physiological changes that are imperceptible to the human eye, guilt or innocence can be established with greater confidence. However, despite the intuitive appeal of measuring autonomic response as a lie detection method, the scientific community is less convinced of the accuracy of this test (see Bashore & Rapp, 1993 for summary).

       Beyond the law enforcement community, several fields depend on discerning truth from fiction. Neuropsychology is an area in clinical psychology facing a growing number of false complaints, specifically in the evaluation of mild head injury cases. Neuropsychologists frequently evaluate cognitive functions such as memory, attention, executive function, etc. to document brain injury. Often in cases of mild head injury there is no discernable brain damage (as documented by advances in neuroimaging such as MRI and CT scans); however the patient complains of an array of neurological and psychological problems (Snow & Hooper, 1994). Individuals with documented brain injury may legitimately receive compensation as a way to cope with their reduced capacity. The incentive of money or services after being diagnosed with a brain injury through a neuropsychological evaluation increases the likelihood that individuals who have sustained an accident may try to feign cognitive problems or exaggerate their deficits. Thus, in addition to evaluating different cognitive functions with standard neuropsychological tests, an accurate assessment of cognitive function should employ methods to test the honesty of the test-taker's answers.

       Special tests have been developed that attempt to detect individuals who are pretending to have cognitive deficits as a result of a brain injury. Many standard neuropsychological tests do not adequately detect malingered neurocognitive deficit (MNCD), because they were designed to measure brain impairment rather than the level of effort or honesty of the test-taker. Neuropsychological tests typically have a low floor and a low ceiling in order to capture the large decreases in cognitive function resulting from severe brain impairment after insults such as traumatic brain injury (TBI), stroke, etc. Individuals with actual brain damage will typically perform poorly on some standard neuropsychological tests. Additionally, someone who is performing poorly on the test in

an attempt to look brain injured will likely score low and have test results similar to someone with real brain impairment. Even if a feigning evaluee has inconsistent results on standard neuropsychological testing, previous research suggests that variable performances on these tests alone may be inadequate indicators of malingering (Heaton, Smith, Lehman, & Vogt, 1978). Using only standard tests of cognitive function does not appear to be effective in detecting individuals malingering neurocognitive deficit. Thus, in cases where the neuropsychologist is attempting to differentiate between real and feigned cognitive deficit, separate tests of malingering increase the probability of detecting individuals who are pretending to have some type of brain damage.

The most common malingering tests measure recognition memory through paper-and-pencil or computer-based formats. Most current malingering instruments are symptom validity tests (SVTs) in which test-takers are presented with a forced-choice format where they choose between a target (previously memorized) and foil (new item) as to which they remember from prior exposure. Results are used to classify someone as honest or malingering. Generally, a number or percentage of correct answers is used to identify someone who is misrepresenting his/her level of cognitive impairment. These recognition tests are relatively easy for even severely brain injured patients, so that when a person with a questionable head injury performs poorly, it suggests the test-taker may be trying to malinger. While the standard malingering tests have a great deal of literature to support their use to detect malingering (Inman and Berry, 2002; Vickery, Berry, Inman, Harris, & Orey, 2001), developing new methods to detect someone feigning cognitive deficits can help identify malingerers who may be able to successfully modify their number of correct answers to reflect the responses of a person with brain damage. In addition, developing new malingering tests using different tasks would be useful to identify individuals who may be aware of existing tests.

Although most of the current malingering tests use an accuracy score from a recognition test to detect people intentionally performing poorly on a cognitive evaluation, there are other methods that can be employed for this purpose. Reaction time (RT) is a test variable that can easily be obtained from a cognitive task. There has been some recent research supporting the utility of RT as a way to detect MNCD. For the most part, malingerers seem to have longer RTs compared to honest responders and responders with a history of a TBI. RT is not only a relatively new method to detect malingerers, but this variable may be more difficult for someone to fake believably since RT should occur automatically when responding honestly.

In addition to using RT as a new way of detecting MNCD, employing a method which would measure what is actually occurring in someone's brain could be even more successful as a malingering detection strategy. Documenting brain activity may be a way to detect malingerers who are able to successfully manipulate their behavioral data (e.g., accuracy and RT) to appear cognitively impaired. This method of detection has the potential advantage of being able to identify someone who is able to "fool" a paper and pencil malingering test. Obtaining a measure of brain activity from individuals should help identify a feigner who knows how to perform on a malingering test. A malingering test method that measures brain activity has the benefit of measuring what is occurring in the brain (e.g., malingering), and therefore potentially identifying people who would evade detection if only traditional detection methods were employed.

In order to understand the potential benefit of using measures of RT and ERP to detect MNCD, a review of the existing literature on standard, cognitive malingering tests is necessary. A summary of the existing studies that have used RT to detect malingering

will follow to illustrate the possible benefit of exploring this newer method of detection. In addition, to understand the utility of using a direct method of recording brain activity to detect MNCD a general overview of EEG/ERP methodology will be described. The few studies that have used ERP to detect feigned cognitive deficits will also be summarized. Finally, overviews will be given of two relevant cognitive phenomena with well-documented RT and ERP patterns. These summaries will illustrate the possibility of detecting MNCD using RT and ERP measures. These new methods may increase the hit rate of detecting malingerers due to the greater difficulty of feigning RT and/or ERP.

*Malingering*

While neuropsychologists devote a great deal of effort to testing general cognitive functions in people, other behaviors are also examined during an evaluation. In the field of psychology, lying for some type of external reward is important enough to address specifically. Malingering is a type of condition classified in the DSM-IV-TR (APA, 2000, p. 739) as a "V-code", or "condition that warrants clinical attention" as opposed to a disorder. It is the conscious production of false symptoms (such as mental illness or cognitive impairment) for an external incentive (such as money). One familiar example of a malingerer would be the character "Corporal Klinger" from the popular 80's TV series, M.A.S.H. He continually dressed in women's clothing in order to get out of military duty in Korea by trying to convince the Army he was a "cross dresser" (which was a dischargeable "disorder" in the military at that time). While this malingerer was portrayed as an amusing and harmless character, in reality people who lie about problems for external incentives can pose a real problem for society. If undetected, malingerers increase the financial burden on society and reduce resources for people who are truly in need.

In the field of neuropsychology, the issue of malingering is a common and costly one. One survey of neuropsychologists revealed estimates that 40% of the people with mild head injury seen for neuropsychological evaluation in litigating circumstances are feigning or exaggerating their cognitive deficits (Mittenberg, Patton, Canyock, & Condit, 2002). The presence of monetary compensation increases the number of evaluees who may be motivated to fake or exaggerate their cognitive problems for a financial gain. Neuropsychologists rely on different ways to detect such individuals. Behavioral checklists have been developed, but more experimental research is needed to confirm their accuracy as a detection strategy (Inman & Berry, 2002). Some standard neuropsychological tests may also be used to detect deviant response patterns. However, as stated earlier, due to the low ceiling for neuropsychological test scores it may be easier for a person pretending to have brain damage to avoid detection on standard neuropsychological tests. There is a need for the development of more tests sensitive to malingering. Most recently, specific, objective measures of malingered neurocognitive deficit (mainly forced-choice procedures testing recognition memory) were developed to increase clinician confidence that poor performance on such a test is most likely due to intentionally missing answers rather than the presence of brain impairment.

*Detecting Malingered Neurocognitive Deficit (MNCD)*

Clinical frameworks for identifying malingering vary. Although no formal criteria are given for arriving at a determination of malingering, guidelines in the DSM-IV-TR (APA, 2000) instruct a clinician to strongly suspect malingering if two or more of the following conditions are present: a medicolegal context, discrepancy between subjective and objective information, lack of cooperation with the assessment or treatment, or the presence of Antisocial Personality Disorder. These criteria have been

severely criticized (Rogers, 1997), and are probably applicable only to psychiatric malingering.  For the diagnosis of malingering cognitive deficits, Slick, Sherman, and Iverson (1999) developed a set of criteria for identifying *Definite* malingered neurocognitive deficits (MNCD) that is more specific than the general guidelines outlined in the DSM-IV-TR.  Those criteria are 1) the presence of a substantial external incentive, 2) definite negative response bias, as demonstrated by significantly worse-than-chance performance on a well-validated measure of malingering, and 3) the behaviors meeting the necessary criteria from the test data are not fully accounted for by psychiatric, neurological, or developmental factors.

Slick and colleagues (1999) also propose definitions and criteria for *Probable* and *Possible* malingered neurocognitive dysfunction.  *Probable* MNCD is documented by the presence of evidence *strongly* suggesting intentional exaggeration or fabrication of cognitive dysfunction without plausible alternative explanations.  The criteria for *Probable* MNCD are 1) the presence of a substantial external incentive, 2) two or more types of evidence of fabrication or exaggeration from neuropsychological testing, excluding definite negative response bias *or* one type of evidence from neuropsychological testing (excluding definite negative response bias) and one or more types of evidence of fabrication or exaggeration from self-report, and 3) the behaviors meeting the necessary criteria from the test data or self-report are not fully accounted for by psychiatric, neurological, or developmental factors.

*Possible* MNCD is indicated by evidence for intentional exaggeration or fabrication of cognitive dysfunction without plausible alternatives.  The criteria for *Possible* MNCD are 1) the presence of a substantial incentive, 2) evidence of fabrication from self-report, and 3) behaviors meeting the criteria from self-report are not fully accounted for by psychiatric, neurological, or developmental factors.  Recent published research has begun to apply the Slick et al. (1999) criteria to the detection of cognitive malingering (Greve, Bianchini, Mathias, Houston, & Crouch, 2003).

Neuropsychological tests assess cognitive functions that may be depressed as a result of brain damage or dysfunction.  However, in addition to integrity of brain functioning, these tests are also sensitive to a variety of potentially confounding factors, such as medication effects, psychiatric conditions and inadequate effort from the test-taker.  Thus, among other concerns, the validity of neuropsychological tests depends on obtaining optimal effort from the test-taker.  In compensation-seeking circumstances, individuals may deliberately under-perform on tests designed to measure their current cognitive ability.  As previously noted, this phenomenon is generally known as malingering neurocognitive deficits.  Individuals involved in circumstances in which there is potential monetary incentive for being impaired, such as in litigation or worker's compensation cases, are thought to be more likely to feign neuropsychological impairment (Pankratz & Binder, 1997).  This concern prompted the creation of separate, objective indices specifically intended to detect feigned cognitive deficits.  These tests were designed to be relatively insensitive to brain dysfunction but broadly reflective of the level of effort given.

Neuropsychological tests are designed to detect and measure significant brain impairment, so they tend to be relatively easy for the unimpaired individual.  Therefore, in order to feign cognitive deficits, the typical intact individual must either deliberately answer the test items incorrectly or exert little effort to perform adequately.  While intentionally faking answers or expending inadequate effort on malingering tests, malingerers may actually go too far and perform more poorly than an individual with

4

significant brain impairment. In extreme cases, malingerers may perform significantly below chance on dichotomous, forced-choice recognition memory tasks, providing evidence for intentional feigning. However, Guilmette, Hart, & Giuiano (1993) reported that use of a strict significantly below chance performance is insufficiently sensitive to more subtle forms of malingering. This dilemma resulted in the current practice of comparing results on malingering tests from compensation-seeking individuals with questionable evidence of brain damage to non-compensation-seeking groups with objective evidence of moderate to severe head injury. Performances on malingering tests falling below cutoff scores established in the non-compensation-seeking, but significantly cognitively impaired group raise the possibility of malingering in individuals with little or no documented brain damage. Determining the relative efficacy of these tests as well as establishing cutoff scores to detect MNCD is achieved through evaluating new malingering tests using a series of methodological designs.

The process by which new malingering tests are "proven" to successfully identify malingerers is important in evaluating and developing new methods of detecting cognitive malingering. Understanding the accepted methodology used in validating new detection strategies for MNCD is especially important when incorporating research tasks testing different types of cognitive phenomena or when using a relatively new data collection technique to detect malingering, such as ERP. The data obtained in malingering studies using a new malingering test method can be interpreted with greater confidence if these malingering research methodological guidelines are followed as closely as possible.

*Methodology for Malingering Research*

Before a malingering test can be used in a neuropsychological evaluation, it must be validated. Three main research designs have been utilized to validate malingering tests and determine their accuracy in classifying responders as honest or malingering. A simulation study design typically uses "unimpaired" or nonclinical populations such as Psychology 100 students and randomly assigns the participants to respond to a test honestly or with instructions to "fake bad" (in other words, to look cognitively or psychiatrically impaired on the tests). The accuracy of the test is estimated with reference to the number of people in each group correctly classified on the basis of their test results. While the simulation study design is appropriate for an initial validation study for a malingering test due to its high internal validity, it has little generalizability to relevant clinical populations unless a "clinical honest" group is included. In the area of neuropsychology, a "clinical honest" group is typically defined as a group of participants with documented brain injury but no known motive to malinger. A major drawback of this design is its lack of external validity in that it does not examine the test's utility in an actual forensic population (real-world malingerers).

Another frequently used research design in the area of malingering is called a differential-prevalence design. This design compares malingering test scores of a group at high risk for feigning (compensation seekers) to a group with low risk for feigning (non-compensation seekers). The malingering test's positive classification rate should be higher in the compensation-seeking group, because there are thought to be more individuals exaggerating in this group. If no difference is found in the rates of positive (malingering) scores in the two groups, serious questions about the validity of the malingering measure should arise. The external validity of this method is higher than in the simulation study, because individuals actually undergoing clinical and forensic evaluations are used as participants. However, a major drawback to this research design

is that it cannot be used to assess the accuracy of test classifications, as the exact criterion status of each subject is unknown.

Due to its logistic difficulty, a known-groups design is generally the last methodology used in malingering test validation. This approach assesses the accuracy of a test's classification rate in actual clinical or forensic samples. A "gold standard" malingering measure is used to assign the evaluees to honest or malingering groups. The accuracy of the new test is established by determining its classification rate in the two groups identified by the criterion measure. This research design has higher external validity than the simulation study because actual clinical populations undergoing evaluations are evaluated. However, a drawback for this design is that the validity of the gold standard test determines the ceiling of the estimated accuracy of the new test. In addition, it is possible that only blatant malingerers may be identified by the external test and thus, the results may be somewhat limited in terms of generalizability. Finally, because true experimental control over the variable of malingering is not achieved, the issue of causality cannot be addressed.

Rogers (1997) advocates requiring multiple supportive findings across the different designs before accepting a malingering test as valid. When examining the potential of a new detection strategy or type of cognitive task to use as a malingering test, simulation studies are viewed as good initial designs because of the high level of control over the subject's response style (assigned malingerers vs. honest responders). Differential-prevalence designs give information about the construct validity of the new procedure in clinical groups thought to have different base rates of feigning. Known-group designs also use actual clinical populations to raise external validity, and this design addresses the problem of internal validity by classifying individual evaluees as honest or malingering based on the test sign from a gold standard malingering measure. Minimally, validation using both simulation studies and known-groups designs is thought to be necessary for establishing the efficacy of a new malingering measure (Berry, Baer, Rinaldo, & Wetter, 2002).

*Effectiveness of Current Neurocognitive Malingering Tests*

As mentioned earlier, methods such as behavioral checklists and response inconsistencies on standard neuropsychological tests have been used to detect malingerers. However, the most common method used to detect MNCD is Symptom Validity Testing (SVT), procedures where forced-choice tests assess a person's recognition memory. With this two-alternative forced-choice task, if an individual scores statistically significantly below chance, it is strong evidence that the test-taker is intentionally choosing the wrong answers, since pure guessing should yield approximately 50% correct answers. As noted earlier, while this evidence is very compelling as an indicator of feigning cognitive impairment, malingering individuals rarely have error rates significantly below chance so cutoff scores were developed to identify malingerers (Guilmette et al., 1993). Ideally, a test should have a sensitivity (percent of those with a condition [e.g., malingering] with a positive test sign) and specificity (percent of those without a condition [e.g., honest] with a negative test sign) as close to 100% as possible where all feigners are labeled as such by their test score falling below the cutoff and all honest responders score above the cutoff. In reality this figure has not been attained. Thus, one advantage to developing a detection method measuring brain activity is that it could have excellent hit rates since physiological measures should establish what is actually occurring in a malingerer's brain (in the case of EEG) rather than his/her performance. However, in the case of current well researched malingering

measures, several forced-choice tests have been examined and have quite high sensitivity and specificity.

A recent meta-analysis examined 5 malingering tests to determine the relative effectiveness of some of the most common malingering tests (Vickery et al., 2001). Effect sizes (ES) are standardized figures that represent the difference between two groups that can be used to compare data from different studies. Cohen (1988) attempted to define guidelines for comparing ES data. In general, a small ES is considered to be Cohen's $d = .2$, a moderate ES is considered to be Cohen's $d = .5$, and a large ES is considered to be Cohen's $d = .8$. ES and classification accuracies were compared for the various tests. The authors conclude that the Digit Memory Test (DMT; Hiscock & Hiscock, 1989) appears to be the malingering test with the overall strongest classification accuracy (83% sensitivity and 95% specificity) as well as large effect size ($d = 1.95$). The Portland Digit Recognition Test (PDRT; Binder & Willis, 1991) had a lower sensitivity than the DMT (44%), but the specificity was excellent (97%). Both the DMT and the PDRT are SVTs which present the test-taker with 2 multi-digit numbers (target and foil number) after a target number is presented. The test-taker then must recognize which of the 2 numbers was shown earlier. While the review supports the research using the DMT and PDRT specifically, the authors noted that a limitation of a number of the studies conducted on the malingering tests reviewed in their meta-analysis was the lack of a clinical control group of participants with brain injury. There is reduced generalizability of the test results when studies use only healthy participants instructed to feign while taking the test. While using a simulation design comparing healthy honest and healthy individuals instructed to feign on test in a natural first step in validating a potentially useful technique to identify individuals with MNCD, there may be a big difference in how the test detects real world fakers compared to honest patients with actual brain impairment.

Inman and Berry (2002) followed methodological recommendations listed by Rogers (1997) in order to examine existing malingering tests using a research design with greater generalizability to real world malingerers. Inman and colleagues incorporated several methodological additions such as the inclusion of a control group of individuals with a history of mild head injury, matching groups on demographic variables, using a battery of tests as in a real world evaluation, and post-test manipulation checks to ensure malingerers and honest responders remembered their role throughout the testing. These changes in the research design also enhance a clinician's confidence in the malingering tests that performed best under these more stringent guidelines which more accurately reflects the forensic evaluation setting. The DMT and the Letter Memory Test (LMT; Inman, Vickery, Berry, Lamb, Edwards & Smith, 1998) had the highest possible specificity (100%) and high sensitivities (64% and 73% for the DMT and LMT respectively). Similar to the DMT, the LMT is a SVT which tests recognition memory for a random string of letters where the forced-choice pair (target and foil) is presented after a studied item. The authors conclude that despite the effectiveness of the DMT and LMT, new tests of malingering need to be developed in order to continue to detect people feigning cognitive deficit. The tests reviewed in these meta-analyses were tests of recognition memory.

Vickery, Berry, Dearth, Vagnini, Baser, Crager, & Orey (2004) examined the LMT, DMT, and the Test of Memory Malingering (TOMM; Tombaugh, 1996) in a group of head injured and healthy control participants with instructions to malinger or answer honestly. They found that while the presence of a head injury did not affect scores on the

malingering tests, malingering instructions were associated with significantly lower malingering test scores. In addition, the TOMM had the highest specificity (100%) of the three malingering tests examined. The TOMM never classified an honest responder (healthy control or head injured) as a feigner. In contrast to the DMT and LMT mentioned above, the TOMM uses line drawing pictures of common objects that are presented initially in a study phase (all 50 pictures), and then forced-choice pairs (target and foil) are presented (50 pairs presented).

Current malingering tests have a variety of formats that employ stimuli such as pictures, numbers or letters. These recognition tests also have varied study phase methods. All items can be studied before a recognition phase (such as the TOMM) or the recognition phase could occur in varying amounts of time (5-30 seconds) immediately after each target is presented (e.g., the DMT, LMT, or PDRT). One important consideration when trying to detect malingering using a new method is to use a task that is a modification of a standard motivational test. If a new task is only a slight modification of a validated test it increases the validity of the new data generated since the format is very similar to an established malingering test. When attempting to detect MNCD using new methods employing a task that resembles a standard malingering test would be ideal. Interestingly, the TOMM's format of study phases followed by test phases of all forced-choice pairs is comparable to a well known task that has been studied extensively in the cognitive literature. As mentioned earlier, the development of new ways to detect MNCD could be useful as confirmatory evidence of feigning. A brief description of the TOMM will follow to illustrate how successful this test is as a standard malingering test, and thus an ideal malingering test to modify in order to examine the usefulness of new detection methods in this area.

*Test of Memory Malingering (TOMM).*

The Test of Memory Malingering (TOMM; Tombaugh, 1996) consists of 50 line drawings of common objects as target stimuli to be remembered. There are three trials. In Trial 1, the 50 pictures are presented one at a time for 3 seconds each. After presentation of the 50 drawings, a recognition testing phase is administered, consisting of 50 trials, each having one of the previously shown target drawings paired with a foil. The test-taker chooses the pictures he or she remembers having been shown previously. In Trial 2, the same format as Trial 1 is used and the study phase of 50 old drawings from Trial 1 is repeated, followed by a test phase involving 50 pairs of pictures (target and new foil pictures) from which the previously seen item must be chosen. A 3rd retention trial may also be administered after a delay of approximately 15 minutes. The retention trial involves only the test phase with the 50 target drawings paired with 50 new foils; the test-taker chooses the one he or she remembers being shown before the delay. Development of this test was based on evidence that recognition memory (as opposed to free recall memory) is typically left intact even after severe neurological impairment. Thus, the TOMM manipulates face difficulty level by presenting a task that is erroneously thought by many test-takers to be difficult.

The development of the TOMM was conducted in two phases using nonclinical subjects and a clinical sample with diverse cognitive impairments (Tombaugh, 1997). The nonclinical population was found to be extremely accurate at identifying targets. In the first trial 94% of the targets were correctly identified as were over 99% of the targets for the remaining two trials. Individuals with various cognitive impairments also identified targets at a rate almost as high as the nonclinical population, establishing the test's insensitivity to significant brain impairment. Based on these results, a cutoff of

45/50 (90%) was established where scores below 45/50 for a given trial suggest poor effort.

In the second series of studies, the test was first validated on a group of "at-risk" malingerers in a differential-prevalence design. The scores for the compensation-seeking group were significantly lower than for the previous two groups of honest responders. Additionally, in a simulation study, a sensitivity and specificity of 100% for individuals instructed to fake or respond honestly on the TOMM were obtained with the suggested cutoff score. Finally a criterion of 45/50 (90% correct) on the second trial correctly classified 95% of all non-demented patients and 91% of all patients (including patients with Alzheimer's Disease) as not malingering in a validation sample of 475 community volunteers and 161 neurologically impaired patients (Tombaugh, 1997). A later study by Rees, Tombaugh, Gansler, and Moczynski (1998) demonstrated converging validity using simulators, high-risk populations for malingering to occur, and a computer form of the test. Reliability coefficients were not reported for the TOMM. Based on this evidence, the TOMM appears to be a well-validated motivational test.

*New Methods to Detect MNCD*

New methods of detecting malingering are essential in order to continue to identify individuals who may be coached on how to respond to existing tests. The TOMM is of particular interest for present purposes, because it is very similar to a task which measures recognition using accuracy and RT. A task that closely resembles a traditional malingering test with a consistent RT pattern (e.g., different RTs to different items) and an established pattern of brain activity (i.e., different brain responses to different items) would be an ideal way to test the role of RT and ERP in detecting MNCD. A slight modification of the TOMM would make it more compatible with different malingering detection methods. By presenting one stimulus at a time instead of target and foil pairs, there would be very little difference from the original forced-choice format. This modification would allow for RTs to be obtained to different items and ERP recordings for each item. Potentially, interesting comparisons of honest and malingering responses can be made with the new data generated using a modified TOMM task. The presentation of one item at a time (instead of pairs) would allow for data from brain activity to be collected using ERP methods as well as a newer behavioral variable of RT (specifically, patterns of RTs to different items). Before illustrating how this modified task could be used to detect malingering using ERP, a brief review of the effectiveness of recent studies that have used RT to detect MNCD will show the potential benefit of using this newer behavioral method.

*RT and Malingering*

Most of the available malingering tests used to detect feigned neurocognitive deficits use the percentage of correct responses to identify malingering individuals. Reaction time (RT) is another variable that can be easily extracted from computerized malingering tests. Compared to getting a certain number of answers wrong on a task, manipulating RT may be more difficult for would-be malingerers to accurately mimic RTs from brain injured test-takers. vanGorp and colleagues suggest that RT is difficult for feigners to modify believably, particularly on tasks which compare more than one score from a test to obtain a difference score (e.g., Trail Making Test A & B) or a comparison score (Grooved Pegboard right vs. left hand RTs) (vanGorp, Humphrey, Kalechstein, Brumm, McMullen, Stoddard, & Pachana, 1999). Malingering participants did not show a typical difference in their RT when comparing the RT to trace an easy "trail" (Trail Making Test A) to a RT to trace a more difficult trail (Trail Making Test B).

The RT to an easy tracing pattern should be faster than the RT to a more complicated pattern. These individuals also did not show a standard difference in right hand RT compared to left hand RT (the left hand is generally slower than the right hand in right-handed individuals). It is thought that malingerers may not be aware of all of the scores that are being compared in such tasks (they may just be trying to look "slow") and as such, they may not show typical variations in RT that most honest responding individuals produce.

For a measure to be successful in detecting malingering, brain-injured patients should score about as well as intact individuals responding honestly. Typically, even moderately head injured individuals show the same patterns as intact responders (with slower overall RTs) such as those mentioned for Trail Making Tests and Grooved Pegboard (vanGorp et al., 1999). In order to determine the utility of RT as a malingering detection strategy, dissimulating responses should be compared to RTs from a group of brain injured participants in order to ensure that this method would not misclassify a brain injured respondent as malingering.

The ability for individuals to feign cognitive deficits on two information processing tasks (an auditory reaction time test and the Paced Auditory Serial Addition Test (PASAT; Gronwall, 1977) has been examined in order to see how effective using an information processing task to detect malingering would be (Strauss, Spellacy, Hunter, & Berry, 1994). The reaction time task was more effective than the PASAT in classifying the three groups (honest, malingering, and brain injured control groups). The malingering group mean RT (combining right and left hand RTs) ($M = 495$ ms) was significantly slower than the honest ($M = 169$ ms) and participants with a closed head injury (CHI) ($M = 223$ ms) groups. The honest and CHI groups were not significantly different. Using RT to classify responders yielded a 76% hit rate for honest (control and brain injured) and malingering groups. These results suggest that RT appears to have promise as a method of detecting feigned neurocognitive deficit. Although there have only been a few studies examining the effectiveness of RT as a way to differentiate malingerers from honest and brain injured responders, most of the research suggests that malingerers perform significantly slower than both non-brain injured control and brain injured control groups (Strauss et al., 1989; Osimani, Alon, Berger, & Abarbanel, 1997; vanGorp et al., 1999).

Despite these promising results, not all studies have found that RTs were slowest for malingering participants. Rose, Hall, and Szalda-Petree (1995) included a clinical control group of moderate-to-severely brain injured participants in addition to honest and malingering responders in their examination of the effectiveness of RT to identify malingerers using the computerized version of the Portland Digit Recognition Test (PDRT-C). They found that RTs for the brain injured participants were significantly slower ($M = 4.23$s) than for honest ($M = 1.69$s) and coached and uncoached malingerers ($M = 3.21$s and 2.91s, respectively). One difference between the PDRT-C and the tests mentioned above is that the PDRT-C is a forced choice test with no time limit. Brain injured participants may have been more unsure of their response and waited longer before going onto the next item which could lead to longer RTs compared to a time-limited test.

RT is frequently produced in a specific pattern that can be used as a malingering detection strategy. Osimani and colleagues (1997) explored detecting individuals feigning brain injury using a commonly occurring cognitive phenomenon, the "Stroop Effect". The Stroop effect occurs when RT and error percentage are longer and larger in

an incongruent condition (e.g., reading the color of the ink that a different color word is printed in [saying "blue" when the word "red" is printed in blue ink]) compared to a congruent condition (e.g., reading color words printed in the same color as the word name [saying "blue" when the word "blue" is printed in blue ink]). Osimani et al. (1997) found that honest responders and a brain injured group both showed the Stroop effect, while the malingering groups (both uninformed and coached malingering groups) showed an inverted Stroop effect (RT longer for congruent condition compared to incongruent condition and a greater error percentage for the congruent condition compared to the incongruent condition). These results suggest that using a well-established cognitive phenomenon with a RT pattern that is resistant to brain injury may be an effective way to detect individuals feigning neurocognitive deficit.

Although RT has only recently been studied as a method to detect feigned neurocognitive deficit, there appears to be some support for its use. More individuals who are feigning brain damage may be identified as malingering by using a new variable (in this case, RT). One of the advantages of using RT over the traditional accuracy score to detect MNCD is that RT may be more difficult to feign since it typically appears to be a more unconscious process than accuracy responses (thus, malingering may disrupt that natural process and thus reveal feigning). A method that can detect unconscious processes in the brain may also be a way to detect MNCD. Electroencephalography (EEG) measures electrical activity generated by groups of neurons in the brain responding to different items or psychological states. This measure of brain activity may be useful as a new method of identifying individuals who are engaged in a malingering process during a cognitive task.

*ERP-General Description*

As mentioned earlier, neuropsychological tests measure cognitive abilities by inferring brain function through test performance. While these tests are convenient and can yield a great deal of information about how the brain functions, direct physiological measurement of brain activity has advantages in the study of cognitive phenomena. Examples of commonly employed techniques for measuring cognitive function are brain imaging methods in humans (e.g., PET or functional MRI) or single-cell recording directly measuring neural activity in animals. Electroencephalography (EEG) is a technique that directly measures neural activity by recording electrical activity from the scalp. Event-related potentials (ERP) are averaged EEG signals recorded from the scalp that reflect activity underlying a time-locked perceptual, motor, or cognitive event. ERPs have also been linked to some specific psychological processes (Fabiani, Gratton, & Coles, 2000). Each ERP waveform reflects electrical activity generated by groups of neurons firing in the direction of various scalp areas. For the electrical signal to be large enough to be measured, the neurons must be active at the same time and the groups of neurons must be oriented in the same directions (i.e., the dendritic trees and axons are oriented on the same side of a structure). This type of neuronal organization is limited to structures with organized layers of neurons such as cortex, thalamus, and cerebellum. Therefore, ERPs only measure neural activity effectively from those brain areas.

In a typical ERP study, electrical activity is recorded during different events (e.g., recognizing old and new pictures) and/or during different internal states (e.g., attending to stimuli or not attending to stimuli). The electrical activity is typically recorded across several sites on the scalp. Once the electrical activity is recorded, the data must be processed to attain a plot of electrical activity corresponding to each event. The electrical activity is filtered to remove distorting artifacts (such as the large electrical activity

generated by muscle movement due to eye blinking) and compared to a baseline that is typically drawn from the 100ms before a stimulus is presented. This results in a waveform that corresponds to each condition in a study. These waveforms are plots of the changes in voltage across time, and they are compared for differences in ERP mean amplitude or peak latency (see Figure 1 for an example of standard ERP components). The polarity of the wave peaks (negative going or positive going) also helps describe standard components (e.g., P300, a positive going wave that starts at approximately 300ms). The polarity is affected by many factors such as the location of the reference, distance in synaptic gap or type of input (excitatory or inhibitory). It should also be noted that plots of the traditional positive going wave occurs counter-intuitively beneath the x-axis while negative going wave occurs above it. The peak latency of an ERP waveform is the time between stimulus onset (0ms) and the peak of the response of an ERP component (e.g., P300). Scalp distribution pools the ERP responses across many electrode sites in order to show differences in electrical activity across the scalp at a given time. While exact spatial location is poorly detected with ERP, the source of the activation may be determined by modeling.

The amplitude of an ERP component is measured to indicate the strength of the brain response to a cognitive event. The amplitude is measured in millivolts (mV). Wave amplitude is influenced by the number of neurons firing simultaneously in the same direction reaching a specific area of the scalp. For example, if two groups are compared and one group has larger mean amplitude compared to the other, it may mean that the group with the larger amplitude had more neurons firing in response to an event. This can help to identify group differences in cognitive processing. While the electrical activity recorded from the scalp occurs soon after the neurons actually fire in response to an event (creating excellent temporal resolution for this measurement), the spatial resolution is less accurate with ERP since neurons can be firing from different places in the brain. For instance, ERP mean amplitude may be the largest in the frontal area of the scalp, but the actual group of neurons firing may be located in the central area of the brain and firing towards the front of the scalp. Yet this type of documentation of ERP response is important to identify, because it indicates the presence of a group of coordinated neurons working together during a cognitive task. Identifying group differences in this manner suggests a variation in brain response to events. The mean amplitude of the waveform is a way to document the timing and strength of a cognitive phenomena more than a method of localizing where the event occurs in the brain.

Many cognitive processes have been linked to specific waveforms in the ERP literature. The "Old/New effect" refers to the ERP pattern produced when participants are asked to identify targets (previously studied) and novel items, stimuli which are also used in a traditional malingering test, the TOMM. Studied items evoke a larger positive going ERP waveform compared to the waveform for new items (see Figure 2 for ERP waves documenting the "Old/New" effect), specifically in the posterior area of the scalp. A component is a term used to indicate that portions of ERP waveforms vary in responses to experimental manipulations. For instance, the P300 is an ERP component involved in many cognitive processes. The larger the mean amplitude of this wave, the more processing resources are demanded by a specific task (Fabiani et al., 2000). These well-established waveform patterns in various tasks allow for comparisons between different groups (e.g., controls vs. brain injured participants) in order to examine any differences in brain activity between the groups. Examining brain activity while a person is malingering may help to increase knowledge about this unknown cognitive process.

12

Figure 1
Examples of waveform peak characteristics used for ERP data analysis
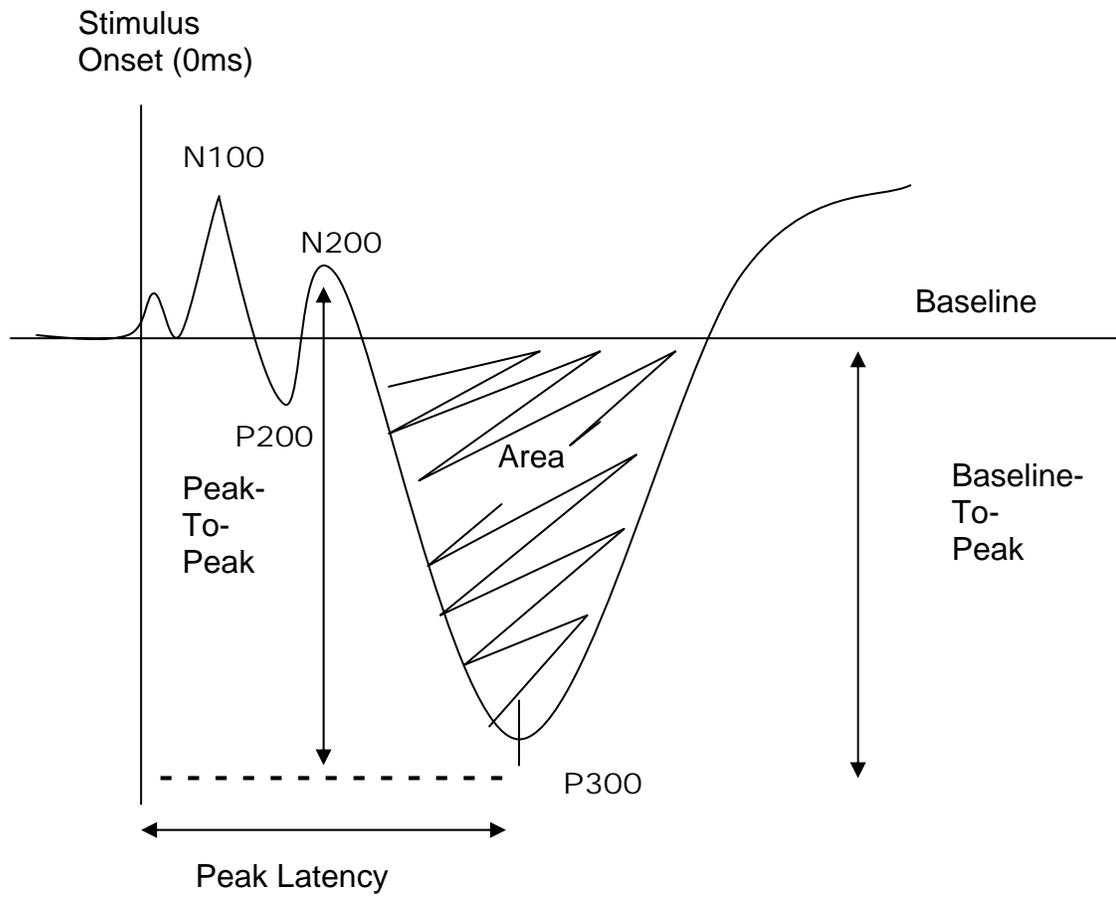


*Note:* Figure adapted from Fabiani et al. (2000) p. 59 in the Handbook of Psychophysiology.

Picton and colleagues (2000) summarized guidelines for conducting and analyzing data using ERP. The authors state that when examining relatively new cognitive phenomena (such as malingering); it is essential to use a well-researched cognitive task to elicit the phenomenon under study. Using a task that has been successful in prior research will help when disseminating data related to the new phenomenon. In addition, when using patient populations in a research study it is essential to match the groups on most/all other characteristics with the exception of the variable being studied. For example, if a group with a history of a head injury is included in a study, groups should be matched as closely as possible on variables of age, gender, education, etc. When analyzing data, all aspects of filtering applied to the ERP data must be noted. In addition, when describing significant effects, ERP waveforms should be presented to aid in the interpretation of the data. The data from a study should also be interpreted in terms of prior research in the area. The guidelines should be followed as closely as possible in order to increase the applicability of ERP results, especially when using ERP to study a relatively new phenomenon like malingering.

*ERP and Malingering.*

As stated above, event-related brain potentials (ERPs) measure brain activity from various places on the scalp, time-locked to specific events. The research using ERPs to detect criminal guilt has employed a procedure called the Guilty Knowledge Test (GKT) where crime relevant items (e.g., the gun used in a hold-up) and crime irrelevant items (e.g., a bow-and-arrow) are shown to the test taker and a specific waveform, the P300, is elicited from "guilty" participants when crime relevant items are shown. This recognition of crime relevant items produces a P300 wave that is not present in people who have no prior exposure to the item. The P300 ERP response occurs when crime relevant (or old) items are recognized by a person. The presence of the wave indicates knowledge of a crime despite claims of innocence by the participant. Currently, the polygraph (a measure of autonomic response) is used more often than ERPs in situations like law enforcement and pre-employment interviews as a deception detection technique. Yet advocates of ERP state that the brain activity is less subjective and may be a more accurate method of detecting deception than the polygraph (Farwell & Donchin, 1991; Bashore & Rapp, 1993).

As mentioned in an earlier section, psychology is a field that depends on accurately assessing an individual's current state. Psychologists evaluate personality and cognitive variables using self-report questionnaires and behavioral tests. These measures are vulnerable to distortion by a person attempting to look more impaired than is actually the case on an evaluation. In the field of neuropsychology, where large monetary settlements can be obtained if a person can show a head injury led to cognitive deficits, malingering tests have been developed to try and detect individuals who are misrepresenting their cognitive function. In the same way that ERPs may be a stronger lie detection strategy than the traditional polygraph, ERPs may also provide another detection method in neuropsychological evaluations. The ability of ERP to document brain activity directly may help identify malingerers who are able to avoid detection on traditional paper-and-pencil tests.

In 1995, the general area of deception-detection using ERP methods was expanded to show that this technique could be applied to detecting malingered neurocognitive deficit. As mentioned above, the most common type of memory that is assessed with malingering tests is recognition memory. In traditional cognitive research, the P300 wave is associated with recognition memory in the sense that old or previously

14

Figure 2
An example of the Old/New effect P300 ERP wave
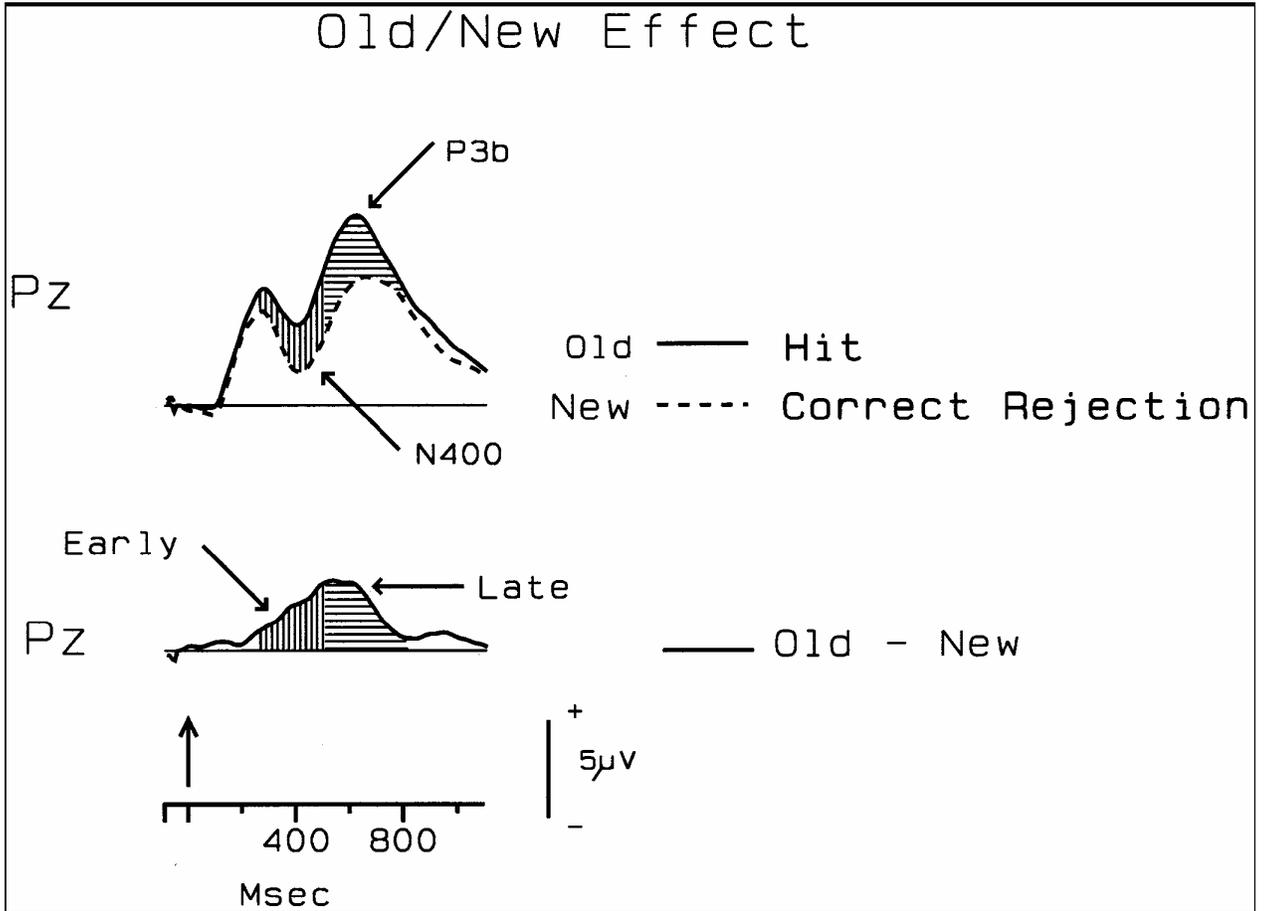


Note: Figure adapted from Friedman (2000) p. 180; positive going wave shown above baseline in this example.

Top row. Grand mean ERPs elicited by previously studied (old) and unstudied (new) words recorded at the midline parietal (Pz) scalp site during the test phase of a study:test paradigm. Bottom row. The result of subtracting the ERPs elicited by new items from the ERPs elicited by old items, i.e. the old:new difference waveform. In the top row, vertical hash marks indicate the P300 region of the old:new effect; in the bottom row, these are referred to as the 'early' and 'late' regions of the old:new effect. Arrow marks stimulus onset, with time lines every 200ms.

seen items produce a larger P300 wave than new items which elicit a much smaller or no P300 wave (Paller & Kutas, 1992). Theoretically, this recognition response would occur if the potential malingerer recognized an item even if the behavioral response was to deny recognition. As previously mentioned, prior literature in deception has made use of this phenomenon as a way to detect lying about committing a crime (e.g., Rosenfeld, Angell, Johnson, & Qin, 1991). In addition to using a recognition task, an oddball task also elicits a P300 wave when infrequent or "oddball" items or information occurs. In simulating amnesia, people may say that they do not remember information such as a birthday or mother's maiden name. In an autobiographical task used in this area of research, the "oddball" is correct autobiographical information about the participant. Individuals should elicit the P300 response to these oddball (accurate) items even if they respond as if they do not remember.

A few studies in the area of malingered neurocognitive deficit (MNCD) have used an oddball task to detect individuals who are malingering. Researchers have examined the effect of feigning memory impairment on the P300 wave produced by an oddball task. The autobiographical knowledge test, which uses the participant's own information (birthday, phone number, mother's maiden name, etc.) as oddball items, was used in a few early studies (e.g., Rosenfeld, Ellwanger, & Sweet, 1995; Ellwanger, Rosenfeld, & Sweet, 1997). This test consisted of a series of 13 items in 3 blocks (a block of dates-one being the participant's birthday, a block of phone numbers, and a block of names). A larger P300 was elicited by the oddball or recognized information (Rosenfeld & Ellwanger, 1999). All honest responders showed this pattern; however, only a few malingerers were detected with a different ERP. This type of oddball task did not yield any false positives (no one honest was classified as feigning), but the sensitivity of detecting malingerers with autobiographical information was believed to be very low because most malingerers knew not to pretend that they do not remember information like their name or birthday. This lack of sensitivity using an autobiographical oddball task prompted the creation of another type of oddball task based more closely on a malingering test.

The next test used with ERP to detect malingering was the P300-enhanced Forced-Choice Procedure (P3FCP), a type of matching-to-sample task. This task was based on the well–validated malingering test developed by Hiscock and Hiscock (1989), an early version of the Digit Memory Test (DMT). The P3FCP presented target and foil items serially and reduced the number of targets (or matches) to 17% or 33% instead of the 50% that typically occurs in the malingering test version when a target and foil pair follows each studied item from which to choose the right answer. The "oddball" items in this case were the reduced number of targets. The largest P300 amplitude was found in the malingering group for rare (or oddball) matches. Rosenfeld, Sweet, Chuang, Ellwanger, & Song (1996) found a 67% hit rate when using ERP (using a larger P300 to identify malingering) in addition to the behavioral test results.

Most of the studies that have used ERP to detect malingering have recorded ERP waves during an oddball task. In examining malingering and honestly responding participants on oddball tasks, some research has suggested that compared to malingering the honest response produces a larger P300 wave. Ellwanger, Rosenfeld, Sweet, and Bhatt (1996) found that although a peak-to-peak analysis (measuring wave amplitude from wave peak to the adjacent wave's peak) showed no difference between the honest and malingering groups, the baseline-to-peak analysis indicated that the honest response produced a more positive P300 wave (see Figure 1 for methods of measuring ERP

16

waveforms). It has been suggested that P300 decreases are due to an increased workload (Johnson, 1986). Thus, it may be that the reduced P300 in the malingering group is due to an increased cognitive demand when trying to malinger responses for a task.

The differences between the honest and malingering responders may also occur in the pattern of brain activity when examining responses to different items. Another study using the P3FCP found that although both the malingering and honest groups had larger P300 waves for matches compared to P300s for mismatch items, the difference was greater when a test-taker was responding honestly (Ellwanger, Tenhula, Rosenfeld, & Sweet, 1999). If this criterion for identifying a responder as honest was set as a significant difference between the P300 for match and mismatched items, malingering detection rates were good 80% sensitivity and 100% specificity. This result is thought to be due to enhanced devotion of recollection processing under honest conditions, and the similarity in the ERPs for matches and mismatches under malingering conditions is thought to be the result of muted brain activity due to competing cognitive demands. Under malingering conditions, the responder not only discriminates between the old and new stimuli but uses additional cognitive processes when evaluating responses in an attempt to look cognitively impaired.

In contrast, other studies found that the malingering group produced a larger P300 wave than the honest group for oddball items during the oddball task, the P3FCP (Rosenfeld, Ellwanger, Nolan, Wu, Bermann, & Sweet, 1999). Rosenfeld, Reinhart, Bhatt, Ellwanger, Gora, Sekera, and Sweet (1998) also found larger P300 amplitude for malingerers. This increase of P300 peak-to-peak amplitude was largest for oddballs (studied items that occurred 1/9 times) and was located in the Pz site. The authors suggest that this greater P300 wave when a test-taker was feigning cognitive deficit was the result of increased cognitive demands as a result of trying to feign performance while answering the test questions. The oddball task has produced mixed results in the few studies that have used this task. Malingerers either produced little difference between ERP for match vs. mismatch items (Rosenfeld et al., 1995; Ellwanger et al., 1996), honest responders produced a larger difference in the P300 wave for match and mismatch items (Ellwanger et al., 1999), or malingerers produced an even larger difference between the ERPs for the two conditions than the honest responders (Rosenfeld et al.,1998; Rosenfeld et al., 1999).

In addition to examining the P300 in malingering paradigms, different scalp topographies were found between honest and malingering groups. Scalp topographies give a general map as to strength of the electrical activity measured at the scalp. For example, a task may elicit the strongest electrical activity from the posterior region that progressively drops off toward the frontal regions of the scalp (see Figure 3 for an example). The scalp distribution of electrical activity may represent another method of differentiating malingering and honest responders. If malingering and honest responding represent two different cognitive states, they should produce different electrophysiological recording distributions on the scalp. Rosenfeld et al. (1999) examined the frontal (Fz), central (Cz), and posterior (Pz) sites on the midline and determined that while honest responders have different scalp distributions for match and mismatch stimuli in a matching-to-sample task, this difference is minimal for responders who are malingering. This may mean that the malingering response becomes the predominant cognitive process beyond recognition. Another study examined additional frontal, central and parietal sites on the left and right side of the scalp (9 sites total) and found that the malingering responders showed Old/New effects earlier, specifically over

17

Figure 3
An example of a topographical map for the Old/New effect comparing electrical
distribution across the scalp for two groups
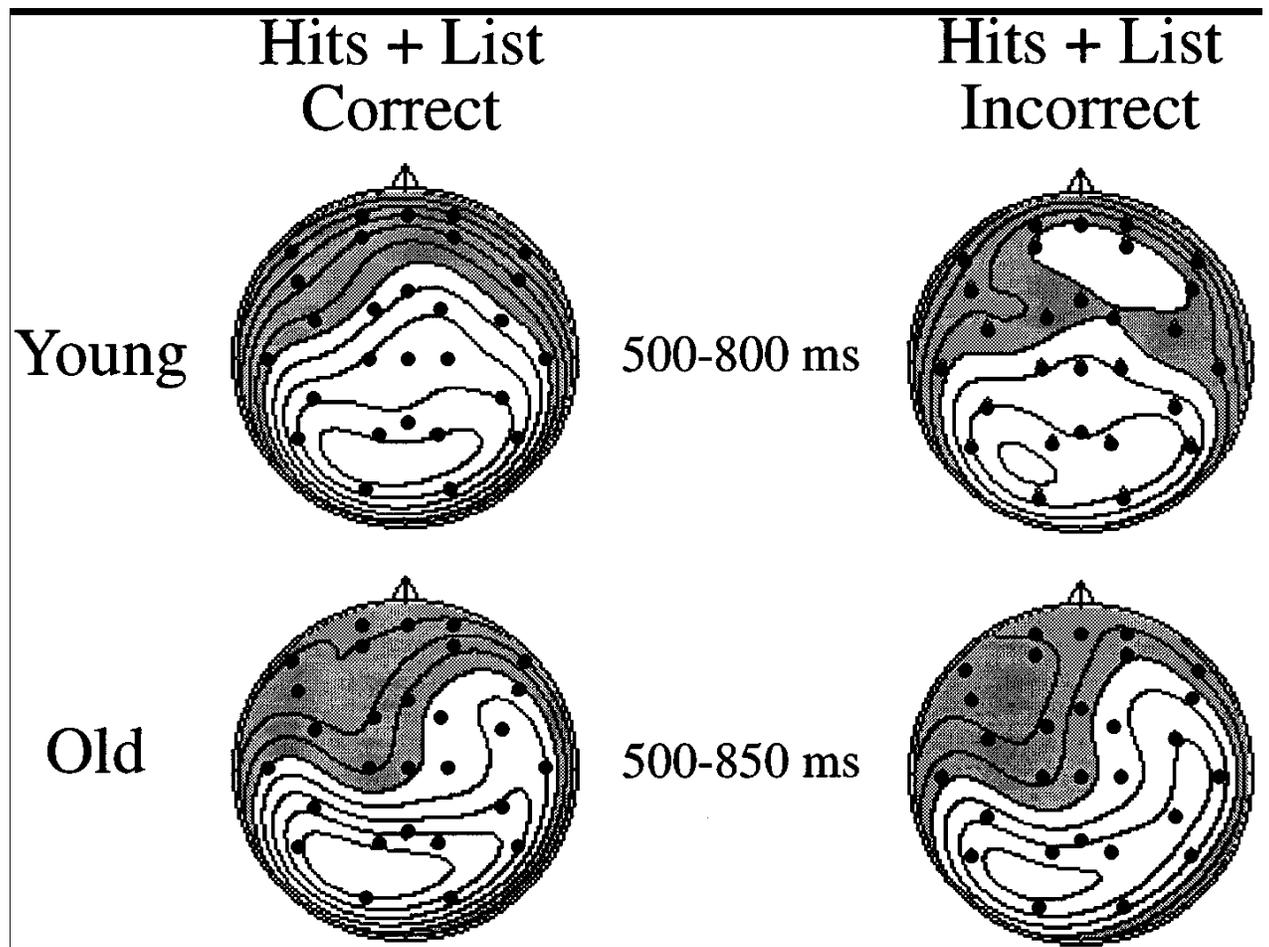


*Note:* Figure from Friedman (2000) p. 197; surface potential scalp maps of the posterior old:new effects
associated with hit trials that were accompanied by correct and incorrect temporal source attributions in
young and old participants. For both top and bottom portions, isopotential lines are separated by 0.40 mV.
Shaded areas represent negativity, unshaded areas positivity.

left and frontal sites (Tardiff, Barry, Fox, & Johnstone, 2000). There was also a borderline significant increase in the difference in peak amplitude between old and new words in malingering participants measured from the left half of the scalp (suggesting a hemisphere effect occurs during malingering). The authors hypothesize that this difference may be related to a greater strategic cognitive functioning occurring in attempting to feign responses.

Most recently, Tardiff, Barry, and Johnstone (2002) instructed participants to respond honestly and under malingering instructions. They found that early on (between 190-320ms after stimulus presentation) the malingering response produced a larger positive wave indicating greater cognitive resources allocated to old items. A topographic analysis was conducted and in the honest condition previous findings were replicated related to the Old/New effect such that the posterior region (Pz) showed more positive amplitude for old items compared to new items (compared to more frontal regions). There was also no topographic distribution of this Old/New or novelty detection effect in the malingering condition. However, there was sustained positivity to old words in the frontal regions as well (i.e., Fz) as well as in the posterior region for the malingering condition. Tardiff et al. (2002) interpret this difference in terms of greater cortical recruitment for working memory processes by the malingerer due to continued evaluation of responses as a way not to make too many mistakes and expose feigning. There appears to be a separate cognitive process that masks the automatic recognition occurring naturally in honest responders.

Tardif and colleagues used an ERP compatible version of the Recognition Memory Test (RMT; Warrington, 1984) where the participant signaled whether the first or second presented word in a pair was a match to a previous target word (Tardif, et al., 2000; Tardif et al., 2002). The authors used RT and ERP measures to obtain an 82% hit rate in detecting malingering. This hit rate is an acceptable starting point when examining the utility of a new detection method, but in order to have clinical relevance a hit rate of 90% or higher is usually desirable. While these results are promising, the Tardiff et al. studies only recorded activity from 9 electrodes over the scalp, and RTs may be distorted since the old and new were both presented after the target (half of the time a target was presented first out of two choices and half of the time a foil was presented first) so the test-taker was not responding to each item independently. Changing the test from the one used by Tardif and colleagues to a task more like a standard malingering task and one that is used in the cognitive literature with ERP methods could increase the interpretability of the ERP and RT results.

*Limitations of Existing Malingering Studies Using ERP.*

There are some limitations in the existing research on using ERP methods to detect MNCD. In melding ERP techniques with malingering research there has been some reduction in the external generalizability of the results. Rogers (1997) describes several important components to the simulation design as a way to validate a potentially useful tool at detecting feigned cognitive deficit. The studies that have used ERP in a malingering paradigm have not incorporated many of these recommendations. For the most part, the research in the area has only compared healthy control subjects answering with instructions to malinger or respond honestly.

As stated in an earlier section, the difference between honest and malingered responses in healthy individuals is an important first step in malingering research, but understanding the effect of brain injury on the method of detecting malingering is essential. Only one study has examined the performance of a head injury control group

19

with ERP.  Ellwanger et al. (1997) compared honest responders and a head injured group on the autobiographical test and a word recognition task where the oddball is expected to elicit a larger P300 wave.  The researchers did not include a group who responded to malingering instructions.  Ellwanger et al. (1997) found that although the P300 amplitude for individuals with head injury was lower than for controls (for oddball and frequent stimuli), the oddball–frequent item amplitude difference score was not significantly different for the two groups.  In other words, this difference score has the potential to be used as an intra-individual analysis to differentiate honest responders (healthy or head injured) from malingerers.  However, these results should be interpreted tentatively since the study did not compare honest, brain injured, and simulating responders.  There are obvious drawbacks to developing methods to detect feigned cognitive deficit without knowing the effect of brain damage on the task.  Without knowing how brain injured individuals perform on a malingering detection test, the risks of misdiagnosing truly impaired people as malingerers are too high.

Another problem with the past research that examined ERP and malingering is the type of test used in the studies.  A test must be compatible with ERP data collection method (i.e., computer based and have serial presentation of stimuli) and must be able to detect malingered responses.  Many of the existing ERP studies have used a variety of tasks that elicit a P300 wave through an oddball task.  An autobiographical knowledge test was used in some of the studies in this area where personal knowledge was presented infrequently (approximately 11%) compared to frequently presented irrelevant information.  This is not a validated malingering test, so the generalizability of the results based on this type of test is low.  Another task that was used presented words where the oddball items were previously learned words (approximately 11%) and the new words made up the 89% of the items.  This is also not a validated malingering test to detect feigned cognitive deficit.  Rosenfeld and others used a task based on a standard malingering test in their ERP design (Rosenfeld, Sweet, Chuang, Ellwanger, & Song, 1996; Rosenfeld et al., 1998; Ellwanger et al., 1999). The test was a modification based on the DMT (Hiscock & Hiscock, 1989).  The problem with this modification was that it greatly changed the format from the traditional malingering test.  While the DMT presents an item to remember and then a target and foil to choose from afterwards, the modification presented one number after the original presentation and used a matching target only 17% or 33% of the time (with 83% or 67% of the presentations being foils, respectively).  While this modification is better than using tests not validated in malingering research, the format of the DMT is altered to a significant extent and very few "oddball" trials are available to create a reliable average ERP wave.

One reason that there may be such contradictory results (such as ERPs for malingerers showing more difference between P300 waves for old and new items in some studies, or almost no difference based on other studies) may be the small number of trials used to make the average waves.  The largest number of trials reported was 24 (Rosenfeld & Ellwanger, 1999) which is very low and decreases confidence that the process being measured is actually malingering vs. honest responding.  Most of the tasks used in this area of research use a version of an oddball test which necessitates having an event that occurs relatively infrequently.  Using a task that has equal numbers of types of items can remedy this drawback.

Recently, malingering research using ERP has modified an existing malingering test more closely.  Tardiff and colleagues used a recognition memory test of the "Old/New effect" by modifying the RMT (Warrington, 1984) where participants respond

to old and new words (Tardif et al., 2000; Tardif et al., 2002).  This modified test yielded enough events for new and old items (50% new, 50% old) to obtain reliable P300 waves.  Yet even with this improvement, there are still problems with the RMT and the way it was modified for use with ERP.  One problem with the RMT is the lack of validation compared to other commonly used malingering tests.  In addition, having participants respond to a serially presented pair (either the first or second word in the pair is the target) distorts the RT.  Rather than responding as fast as possible to each item, the person may wait longer than average for many of the trials to see the second item before answering.  While this presentation is closer to the forced-choice format of malingering tests, the well known data corresponding to the classic Old/New effect (ERP waves, scalp distribution, and RT) are not as relevant when using this modification.  If a malingering test format had a study phase of all pictures and then a test phase (rather than a forced-choice pair after one target item to study is given), and targets and foils are responded to independently instead of as serially presented pairs, it would allow for comparisons with standard Old/New effect data more than with the modifications used by Tardif and colleagues.  Modifying a well-validated malingering test to use in an ERP study that resembles a test eliciting an Old/New effect would integrate the fields of malingering and ERP measures more effectively.

In addition to the shortcomings related to malingering research methodology, there are also problems with the experimental design of studies looking at ERP and malingering in terms of ERP requirements.  These problems reduce the amount of control over the cognitive phenomenon being measured with ERP, in this case, malingering neurocognitive deficit.  The tasks used to elicit a P300 wave were different versions of oddball tasks (11% of the events functioning as oddballs).  A mentioned earlier, these events occurring as oddballs do not yield enough trials to obtain a clear picture of the P300 wave.  Ideally events will yield 30-40 uncompromised trials (no eye-blink or error distortion).  An Old/New task typically has 50% new items and 50% repeated items.  Having as many events as possible increases the probability that 30-40 events will be obtained for an averaged ERP waveform to document the event accurately.  In addition, having these targets and foils presented and responded to independently in the test phase will yield more accurate and useful RTs than the RMT format used by Tardif et al.  The target-foil or foil-target pair presentations where the test-taker responds after seeing the second item in the pair would not allow for the detection of a repetition priming effect.  As stated in an earlier section, there is a well-validated malingering test, the TOMM, where 50% of the items are studied previously and 50% are new items in the test phases.  A modification of the TOMM could include having participants respond to each serially presented item (old or new) independently in order to obtain RT (old/new measure) in addition to ERP data.

*New Tests that Allow for RT and ERP Measures*

Examining the effects of malingering on tasks which yield RT and ERP data could be helpful in clarifying how using these new methods of detecting cognitive malingering will improve identification of feigning test takers.  Comparing individuals responding honestly and trying to feign cognitive impairment using well documented cognitive tasks may be an easier way to distinguish what is occurring during the malingering process.  Any reliable differences between honest and malingering groups in brain waves or scalp distribution of activity from a classic ERP effect could be proposed as a possible marker of malingering or a way to describe the cognitive changes that occur when feigning.  Most of the tasks studied in ERP literature also produce RT to different

test items.  As mentioned earlier, RT is a method of detecting MNCD that has received increased attention as a new way to identify malingering.  Cognitive tasks that produce well-known RT and ERP data from the cognitive literature would be useful methods to detect MNCD.  A brief description of two such tasks will follow.  The first task measures the Old/New effect.  It is not only similar to a standard malingering test, but an ideal test to compare honest and malingering responses using RT and ERP data.  The second task also has a standard pattern of RT and ERP data with a greater unconscious component than the Old/New task which could be useful as a malingering test.

   *Old New Effect Task*

   As mentioned earlier, the TOMM represents a well validated malingering test.  The Old/New task has been well researched in cognitive literature examining conscious memory, and resembles the TOMM in many ways.  The TOMM has a study phase followed by a series of forced choice pairs (containing one target and one foil) and a test-taker chooses which picture he/she remembers.  The Old/New task has a study phase followed by individually presented pictures that the test taker identifies as "memorized" (or old) or "not memorized" (or new).  The slight change in format allows for RT and ERP measures to be obtained.  There are also conceptual similarities to the TOMM and Old/New task.  Most well-validated tests of neurocognitive malingering are forced-choice tests of recognition memory (conscious memory).  The Old/New task is a largely explicit (or conscious) memory test that has two major components, familiarity and recollection (Paller, 2000).  The familiarity is due to the old items which are processed differently than new items-even if the person is not aware the items have already been seen.  The recollection component of the Old/New task occurs when the person consciously identifies items as "old" or "new".  This task has been well researched with established RT and ERP patterns.  The Old/New task format is a close match to a standard forced-choice malingering test (e.g., the TOMM) from a task found in the ERP literature.

   As mentioned in an earlier section, prior research suggests that RT data have promise as another variable to detect malingering (e.g., McGuire & Shores, 1998; Rose et al., 1995).  Adding a RT measure to a modified forced-choice task of recognition memory is a relatively new way to detect MNCD.  As mentioned earlier, the Old/New effect is the term given when studied and new items are identified by the test-taker.  Responding to each item individually allows for a measure of RTs to old (or studied) and new items.  The response to new and studied items has been well researched.  Behavioral data have shown reliably different responses for old (or studied) items compared to new items.  RT data indicate that repeated (old) items are responded to faster than newly presented items (Rugg & Nagy, 1989).  The ability to distinguish studied (old) items from new items depends on the cognitive processes of familiarity and recollection (e.g., Paller, 2000; Curran & Cleary, 2003).  The familiarity of an item leads to old items being responded to more quickly than new items.  Using a task which produces the Old/New effect represents a potential way to test a relatively new method of detecting MNCD using a pattern of RTs.  Deviations from the established pattern of RTs (old RT faster than new RT) are likely to occur when a test taker is malingering.

   The Old/New effect has also been examined in the ERP literature such that brain waves associated with old and new items are well known.  This fact has implications for malingering research, because recording the performance of malingerers responding to old and new items may reveal different ERP responses that could provide documentation of or information about the malingering process.  A serial presentation of stimuli that yields RT data for old and new items also allows for this type of physiological

measurement. The classic "old/new effect" occurs when the ERP to old items have larger mean amplitude than ERP to new items (see Figure 2 for an illustration of this effect using ERP) (Paller, 2001). This effect has both a prefrontal and hippocampal component (Friedman, 2000).

The Old/New effect may also be an excellent task to detect MNCD using ERP, because there are some data indicating this effect is intact even in a population with cognitive deficits (Friedman, 2000). While the old/new effect is intact, depending on the nature of the population being studied, the areas across the scalp show different patterns of activity. Friedman (2000) found reduced frontal activity during an old/new task in elderly participants which he associated with frontal lobe deficits in that population. Wegesin and Nelson (2000) found intact but reduced old/new effects in patients with left temporal lobe epilepsy. There is no research specifically examining the old/new effect in TBI populations; however, there is reason to believe that this type of effect may occur despite brain injury based on prior research with other brain injured populations.

If both healthy and brain injured honest responders show the old/new effect in their ERPs, a person malingering during an Old/New task might be detected if the responses to the stimuli were different from the classic brain waves seen for old and new items. On the other hand, ERPs may reveal that the malingerer has poor behavioral data but shows a classic ERP indicating recognition of old and new items. Physiological data such as ERP represent another avenue to detect MNCD. Data collected based on brain activity in addition to behavioral methods may offer a more direct way to detect malingerers that may be more difficult to believably fake. The Old/New task would allow for the use of RT pattern and ERP data to be used to detect MNCD, using a task that is similar a traditional malingering test.

*Repetition Priming Task*

While the Old/New task is an ideal way to obtain ERP and RT data using a task that is very similar to the TOMM, another commonly used ERP priming task would be potentially useful as another new test to detect malingering. There are two main types of memory tasks, explicit (conscious) and implicit (unconscious). As stated above, the Old/New task is predominantly an explicit memory test, meaning the test taker is aware of what the test measures (number of correctly identified items as old or new). While this task is an excellent modification of a standard malingering test with the ability to measure RT and ERP, adding an additional memory test with measures that are less obvious (since unconscious processes are being measured) to the test taker has advantages in identifying MNCD. An implicit memory test is a measure of unconscious processes such as skill learning (e.g., star tracing task), identifying degraded pictures (e.g., Gollin Figure Task), or priming tasks (e.g., word stem completion tasks). There are several priming task that have been well research using RT and ERP patterns and priming task have also been used recently to try and detect MNCD.

"Priming" is a cognitive phenomenon that can be quantified in many ways. Priming can be operationalized as the number of primed responses (or targets) that were given as answers. Evidence that priming occurred would then be in the form of a greater number of targets compared to novel answers. Word stem completion tests might define priming as the number of previously presented target words that are then used to complete the stem on the later test phase. The greater numbers of targets that have been generated are thought to be an indication of more priming occurring. As the priming test score rises, more priming is thought to occur. On the other hand, priming might also be defined as a faster response to a prime (shorter response latency), and in such cases lower

numbers indicate more priming has occurred.

The first controlled study to test the ability of a priming task to detect simulators of cognitive deficits compared a healthy honest group with memory disordered patients and participants instructed to feign memory impairment (Wiggins and Brandt, 1988). The priming task used in the study was a word stem completion test where priming is measured by the percentage of word stems that are completed by previously seen target words. Simulators had significantly fewer target responses to word stems (45%) compared to memory disordered patients (55%) and honest controls (56%). Despite the apparent promise of this type of perceptual priming task, there are only a handful of subsequent studies examining the use of a priming task to detect MNCD. In addition, there is little continuity between the studies from which to derive any definitive conclusions about the efficacy of priming tests as a tool to detect malingering.

Priming has some research supporting it as a malingering detection strategy. In general, the overall results from the eight available studies (two are unpublished dissertations) suggest a moderate effect size ($d = .8$) for priming tests used to distinguish between honest (healthy and/or brain injured) and malingering responders. There appears to be some support for the ability of a priming task to differentiate between honest and malingering responders; however the mixed results in the few studies attempting to use a priming task do not point to a specific priming task to use. Most of the tasks that were used were word or fragment completion. A picture-based, priming test might be a better task that will ensure brain injured test-takers score as well as honest controls since pictures reduce the negative effect of language difficulties such as aphasias that could impair test performance.

Repetition priming has been well researched in the cognitive literature. While there are no specific studies that have examined repetition priming using ERP with TBI participants, the existing behavior data indicate this type of memory may be resistant to brain injury. Repetition priming has been well researched using more physiological research methods. Many ERP studies have reported differences in brain activity when subjects are exposed to 'novel' (not previously presented) vs. 'repeated' (previously presented) stimuli in memory tasks (e.g., Rugg and Nagy, 1989; Potter, Pickles, Roberts, & Rugg, 1992; Paller, 2001 for a review). Clear differences in brain activation are apparent between 300 and 800ms (P300) following the first and second presentations of a stimulus, reflecting adaptation to the novelty of the stimulus materials. In other words, the P300 for the repetition of the item is larger and occurs earlier than the P300 wave for the initial presentation item.

Repetition priming tasks have been shown to reliably produce different responses for repeated items in behavioral data, fMRI, and ERP studies (Wagner & Koutstaal, 2002). When a stimulus is repeated the rate of neuronal firing is reduced (Brown & Aggleton, 2001). This neural reduction is thought to occur as a result of an increased efficiency in processing the repeated item (Desimone, 1996). Fewer neurons fire for repeated items, but the neurons that are firing synchronize, and produce larger ERP mean amplitude. As stated earlier, this "repetition effect" produces a reliably larger ERP wave (specifically, the P300 component) for repeated items (see Figure 4). In addition, this effect occurs even if the participant is actively engaged in a different cognitive task (i.e., not asked about the repetition of the items or if the item was seen previously).

A repetition priming task has not been used in any published studies as a malingering detection strategy. Yet, as mentioned earlier, no other priming task has conclusive support for its use. Therefore, employing this new repetition priming task

24

Figure 4
A diagram of the 32 channel EEG cap (electrode placement)
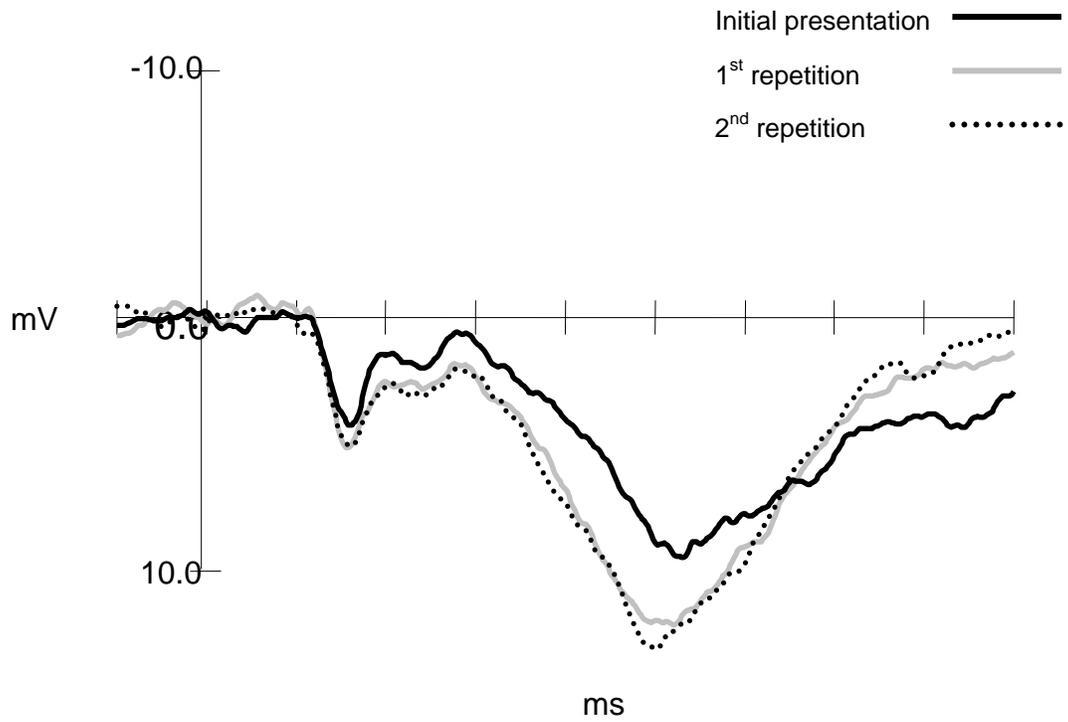
could be more effective in detecting MNCD. The unconscious nature of this type of task may be well-suited to be a malingering test. If a person is trying to appear brain-damaged on a task, he or she should not be aware of the repetition effect and focusing instead on the task itself rather than responding to items as fast as possible. A repetition priming task reliably yields both RT and ERP patterns that can be compared between honest, malingering, and brain-injured responders in order to see if a malingering response can be determined. The Old/New effect task is a necessary first step to seeing the potential for a priming task to detect malingering while applying RT and ERP measures, because it is only a slight modification of a traditional malingering test. However, a pure priming task that measures the effect of repeating stimuli without the conscious recognition of the test-taker should expand the understanding of how RT and ERP can aid in the detection of MNCD.

*Statement of the Problem*

The TOMM is an ideal standard malingering test to use within an ERP framework. With minimal changes, this validated malingering test is similar to a well-research cognitive task called the Old/New effect task. Using a close adaptation of a well-validated malingering test with ERP increases the external relevance of the data obtained and allows the ERP results to be compared to previous research in malingering using paper-and-pencil tests. An Old/New task would present targets and foils individually instead of as forced-choice pairs seen in the traditional malingering test format. This modification of the TOMM is necessary for ERP recordings and RTs to be obtained for different items in the task. The Old/New task was used in this research paradigm to examine the effectiveness of ERP and RT to detect MNCD.

In addition to the Old/New task, a repetition priming task is a priming test that unlike the Old/New task has no recognition component, and it might also be able to identify malingerers due to the unconscious nature of the processes involved. The potential for priming to be an effective malingering test warrants the examination of a true priming test which also has a documented ERP component and RT pattern. A repetition priming task could be a successful test to use in order to see how a priming test performs as a way to detect MNCD without any traditional recognition component to the test. The Old/New and Repetition Priming tasks represent the 2 major types of memory (conscious and unconscious), and in addition to their potential to detect MNCD individually, patterns of responses seen in these different tasks may bolster conclusions drawn about malingering processes.

In short, an Old/New task is a first step in examining the potential for priming to detect MNCD while staying true for the most part to a traditional malingering test format (recognition test). A priming task producing a repetition effect measures priming without a recognition component, and thus this task would be a good second test to see how priming could help detect malingering. As an additional benefit, both of these tasks have been well-researched in the cognitive literature and have documented RT and ERP patterns that could be applied to help identify malingerers from honest and brain-injured responders.

Chapter Two
Method

*Participants*

        Three groups of participants were recruited; a group of normal volunteers with no history of head injury instructed to perform honestly (HON), a group of normal volunteers with no history of head injury instructed to malinger (MAL), and a clinical group of participants with documented brain injury instructed to perform honestly (TBI). Lucas (1998) gave criteria to differentiate between mild, moderate and severe TBI which was followed to create a TBI group moderate to severe TBI for this study.  The criteria that were used in addition to documented brain injury through CT or MRI scans were those that were able to be quantifiable by hospital or EMT professionals on a medical chart.  The criteria to classify a person as having at least a moderate TBI were that the person had lost consciousness $\geq$ 30 minutes and obtained a score of $\leq$ 12 out of 15 on a clinical test that rates verbal responses, eye-opening behavior, and best motor responses called the Glasgow Coma Scale (GCS; Teasdale and Jennett, 1974) which is usually obtained in the emergency room.  Sixteen participants were initially in each group (48 total participants).  After medical records were obtained, one TBI participant was dropped from all analyses because the head injury would be classified as "mild", leaving 15 participants in the TBI group.  In addition, data were missing for one participant from the malingering group for the repetition priming task, because the participant accidentally pressed the wrong buttons for the task (ERP analysis using all responses was obtained for this participant since there is no distinction between right and wrong answers in the analysis).  Normal volunteers were recruited through fliers at the University of Kentucky and Introductory Psychology subject pool classes.  Participants in the TBI group were recruited through fliers at a local private practice neuropsychologist's office and an advertisement in a local newspaper.  TBI participants were required to have sustained their brain injury at least 2 years prior to the study, be their own legal guardian, score $\geq$ 24 on a cognitive screening measure, the Mini Mental Status Exam (MMSE; Folestein, Folestein, & McHugh, 1975), and agree to sign a release for their medical records to be reviewed in order to take part in the study.  In addition, eligible participants were required to have been hospitalized after their injury, experienced a loss of consciousness (LOC) $\geq$ 30 min, and not be currently seeking compensation for their brain injury.  Only right-handed participants were used.  All participants were compensated $10/hour for their time.  TBI participants were compensated an extra $20 as a travel stipend as they were being recruited outside the greater Lexington area.

        The MMSE was used to verify that the cognitive function of the TBI participants was high enough to consent to participate ($\geq$ 24 out of a possible 30).  The mean MMSE score for TBI was 27.67 (*SD* = 1.35) indicating that the group was performing at an acceptable cognitive level to provide consent.  Medical records were obtained for 10 of the participants; medical professionals confirmed the presence of at least moderate brain injury (but provided no specific documentation of the injury such as CT scan, etc.) for three of the remaining TBI participants, and two participants' medical records were unable to be obtained from the hospital but they verbally confirmed the presence of at least a moderate head injury.  For the participants whose records were obtained, the mean GCS was 8.7 (*SD* = 2.94), participants were positive for LOC with a mean duration of 7.21 days (*SD* = 11.98) (however, there was a large range from 30 minutes to approximately 4 weeks), and the mean length of time since the brain injury was 13.18 years (*SD* = 7.21) (ranging from 2 years to 27 years).  The nature of the injury for the

group of TBI participants was primarily motor vehicle accident (MVA) involving mainly cars, one motorcycle and one ATV (73.3%), falls (13.3%), the remaining two participants were injured through damage to the frontal lobe during sinus surgery (6.7%) and a bicycle accident (6.7%).  CT and MRI scans were obtained for many of the participants and the evidence indicated brain injury in varied regions from brain stem, frontal, temporal, occipital and parietal injuries in both the right and left hemisphere.

*Materials*

All participants were hooked up with a 32-channel electrode EEG cap (see Figure 5) using a Neuroscan system.  Participants performed 3 tasks; a computerized version of the Test of Memory Malingering (TOMM-C), an Old/New task, and a repetition priming task.  In addition, all participants completed a brief demographic questionnaire, a Beck-II depression inventory, and a Wechsler Test of Adult Reading (WTAR).

*Test of Memory Malingering-Computerized Version (TOMM-C)*

For the computerized TOMM task, the standard computer format was used with line drawings presented on a computer screen.  Stimulus pictures were displayed on a computer screen approximately 65cm from the participants.  Object size was approximately 8 cm by 10 cm, displayed in front of a white background with a black border.  The computerized version of the TOMM (TOMM-C) included three trials (see Figure 6).  The first 2 trials both have a study and test phase.  For the Trial 1 study phase, 50 line drawing pictures were presented for 3 seconds each.  The participant passively viewed & memorized the pictures.  For the test phase for Trial 1, 50 pairs of pictures (1 target & 1 foil) were presented, and the participant chose the right or left picture (with L-arrow or R-arrow keys) and pressed the "Enter" key to confirm the choice.  For Trial 2, the same 50 pictures were again presented to the participant for 3 seconds each.  This study phase was again followed by a test phase where 50 pairs were presented and the participant chose between the target and a novel foil.  After a 15-minute delay (during which two research assistants began putting on the EEG cap for the next 2 tasks during this delay), the Retention Trial was administered.  Trial 3 (the Retention Trial) consisted only of a test phase where 50 forced-choice pairs were administered (no study phase) and again each pair contained a target and a novel foil.  The program gave feedback to the participant regarding accuracy for all 3 trials.

*Old/New Effect Task*

The Old/New effect is based on previous findings that ERP waveforms to correctly identified old and new items are consistently different.  Generally, in the posterior region of the scalp at approximately 300-700ms after presentation of a stimulus, ERPs to old items are significantly larger than ERPs to novel items.

*Study phase.*

After the TOMM-C Retention Trial was completed, the rest of the EEG cap was put on the participant.  After all electrodes in the cap were connected to the scalp below 10 kilohms impedance, the participant went back into the computer room and viewed 100 new line drawings (Snodgrass & Vanderwart, 1980, & Cycowicz, Friedman, Rothstein, & Snodgrass, 1997) as the study phase for the next task.  These 100 pictures were presented for 5 seconds each, and after a short break, all 100 pictures were shown again.  Stimulus pictures were displayed on a computer screen approximately 65cm from the participants with a visual angle of approximately 7 degrees.  Object size was approximately 8 cm by 6 cm, displayed in front of a black background.  Later, 60 of the memorized pictures in this study phase for the Old/New effect task were used in the last task (Repetition Priming task) as "old" items in order to measure priming of these previously seen items during a

Figure 5
An example of ERP waves representing the "repetition effect"



Electrodes around the eyes

ABREVIATION KEY
REF = reference
GRO = ground
F = frontal area
C = central
T = temporal
P = parietal
O = occipital
M = mastoid reference
HEO = horizontal electro-oculogram
VEO = vertical electro-oculogram
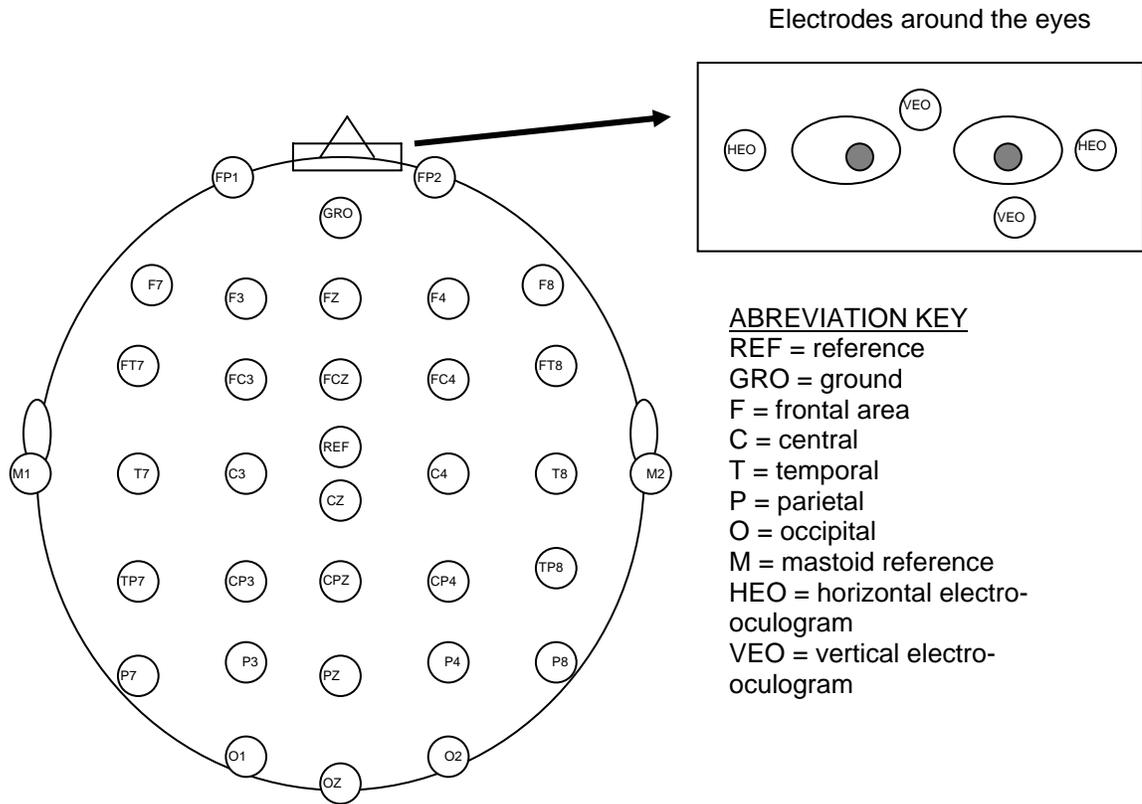
Figure 6
An illustration of the Test of Memory Malingering-Computerized (TOMM-C)

## **TOMM-Trial 1**

Study phase (50

Test phase (50 forced-choice pairs)

3s

3s

3s

...

"correct"

"incorrect"

...

## **TOMM-Trial 2**

3s

3s

3s

...

"correct"

"incorrect"

...
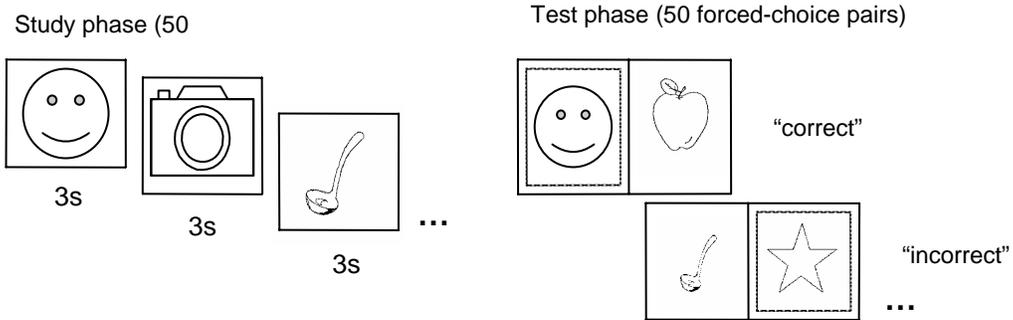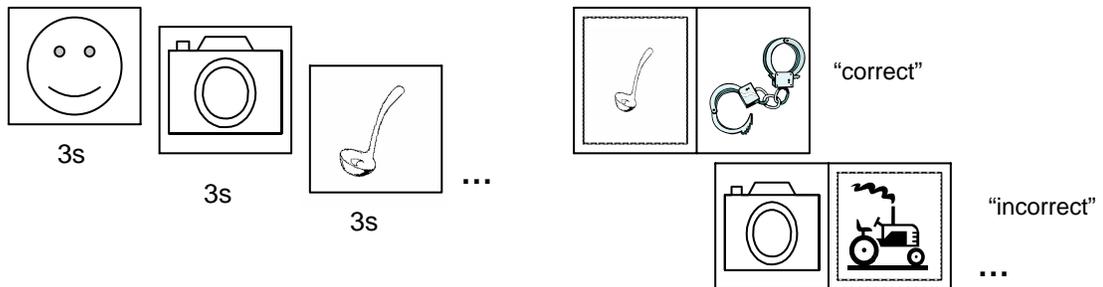
## **TOMM-Retention Trial**

Test phase (50 forced-choice pairs) - no study phase

"correct"

"incorrect"

...
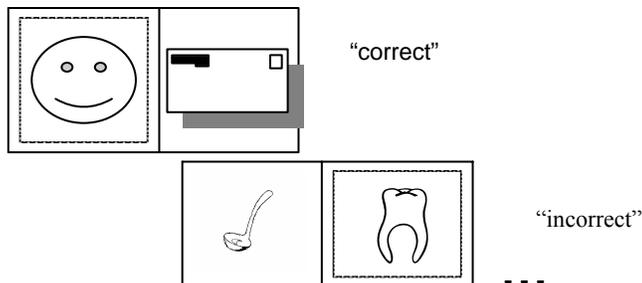
task where the participant was asked to classify objects as "man made" or "not man made" as opposed to noting whether the pictures were old or new (obscuring the nature of one of the comparisons of interest, old vs. new items).

*Test Phase.*

After studying the 100 pictures for the Old/New Effect task, the participants viewed 140 pictures, presented one at a time (70 old pictures previously memorized during the study phase, and 70 foils that had never been seen by the participant) (see Figure 7). For each picture, the participant decided whether the drawing is "new" (for example, "a" key on keyboard labeled as "new") or "old" (for example, "l" key on keyboard labeled as "old") and picked the corresponding key on the keyboard. Stimulus onset was delayed randomly at 100ms, 300ms, or 500ms in order to avoid an expectation effect of exactly when the pictures would appear on the screen. Each picture was presented for 1000ms with an inter-stimulus interval (ISI) of 1100-1500 ms. Fixations were presented for 1500 ms. Participants stared at the screen without blinking until a fixation cross ( + ) appeared, indicating that they could blink. There was one trial of this task lasting approximately 9.5 minutes.

*Repetition Priming Task*

Repetition priming is a type of implicit memory that occurs without the conscious effort of the test taker. This unconscious component is different than forced-choice recognition or old/new effect tasks, because the priming response takes place even while the participant is performing a competing task (such as classifying objects). When responses are faster to previously studied items compared to novel items priming is thought to occur. When RT for a repetition of an item is faster than the initial presentation priming is also thought to occur. In addition, ERP waveforms show larger waves for old items compared to new items (old/new effect) and for repeated items compared to the initial presentation of an item. Since the dependant variables measured in the repetition priming task (old vs. new, and 1st presentation vs. 2nd presentation vs. 3rd presentation) were not the variables that the test-taker answered (classification of a picture as "man-made" or "not man-made"), this task may be more difficult than recognition tests for would-be malingerers to feign believably through their behavioral or ERP responses.

For the repetition priming task, the participants viewed 120 line-drawing pictures (Snodgrass & Vanderwart, 1980, & Cycowicz et al., 1997) presented one at a time (60 old pictures previously memorized during the study phase for the Old/New task, and 60 novel pictures) (see Figure 8). Stimulus pictures were displayed on a computer screen approximately 65cm from the participants with a visual angle of approximately 7 degrees. Object size was approximately 8 cm by 6 cm, displayed in front of a black background.

Twenty new pictures and 20 old pictures (from the 100 previously memorized pictures from the old/new effect task) were presented three times (for a total of 120 mages presented) in the first block of pictures for the repetition priming task. A second block followed with 20 more new pictures and 20 more old pictures, and again each picture was presented three times (for a total of 120 images presented to the participant). Lastly, a third trial block followed with 20 more new pictures and 20 more old pictures, and again each picture was presented 3 times (for a total of 120 images presented to the participant). The participant decided whether a picture on the screen was "man-made" or "not man-made" and picked the corresponding key on the keyboard (for example, "a" key on keyboard labeled as "man-made" or "l" key on keyboard labeled as "not man-made").

Figure 7
An illustration of the Old/New task

Study Phase (100 pictures)
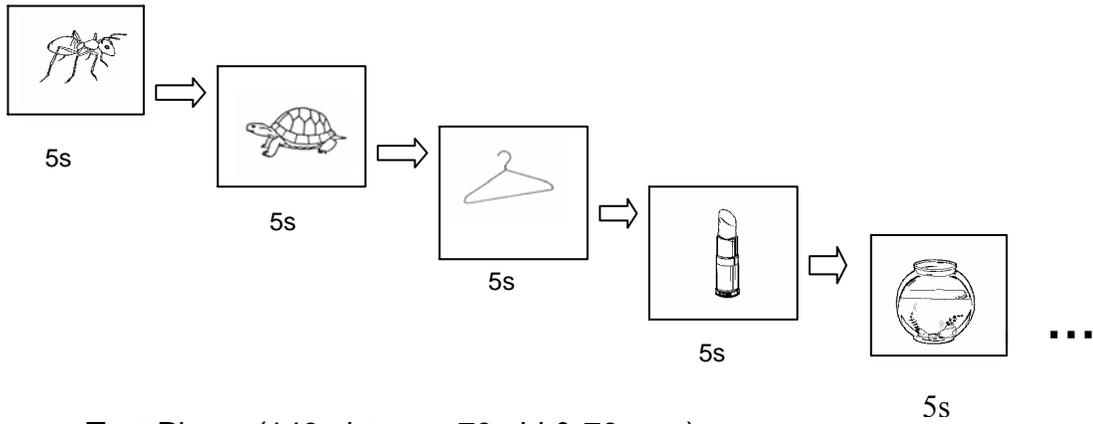


Test Phase (140 pictures: 70 old & 70 new)
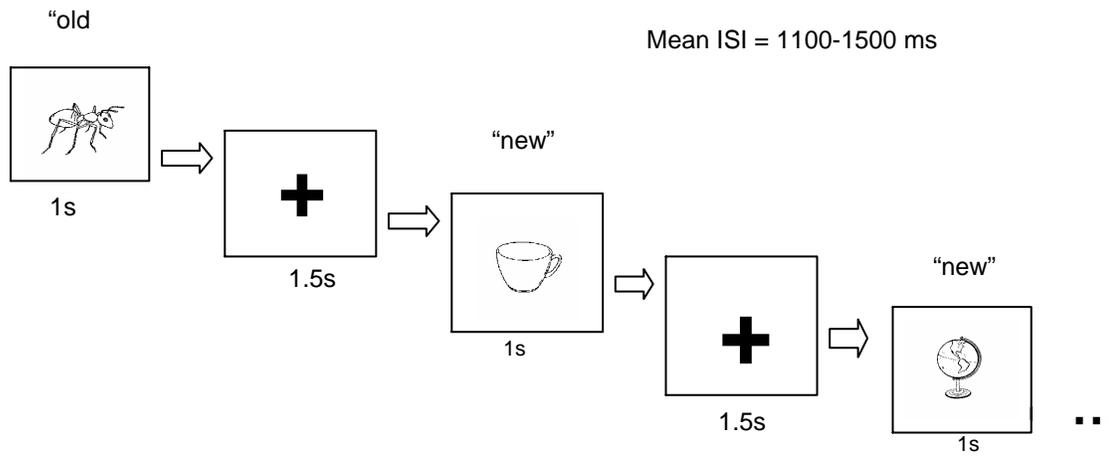
Figure 8: An illustration of the Repetition Priming task

**Trial**

Not man-made, 1st presentation (old)

20 old pictures

20 new pictures

Mean ISI = 1100-1500 ms

1.5s

1.5s

Man-made, 1st presentation (new)

1.5s

1.5s

Not man-made, 2nd presentation (old)

1.5s

...

1.5s

**Trial**

20 old pictures

20 new pictures

Man-made, 1st presentation (new)

Man-made, 2nd presentation (new)

1.5s

1.5s

Not man-made, 1st presentation (new)

1.5s

Not man-made, 2nd presentation (old)

1.5s

...

1.5s

**Trial**

20 old pictures

20 new pictures

Not man-made, 1st presentation (old)

1.5s

Man-made, 1st presentation (old)

1.5s

1.5s

Not man-made, 1st presentation (new)

1.5s

...

1.5s

33

Pictures and fixation crosses remained on the screen for 1.5 seconds, and fixation crosses ( + ) were presented randomly 1/3 of the time on the screen indicating that the participant could blink. The three blocks of this task lasted approximately 7-7.5 minutes each.

Two types of tasks are generally used to elicit priming. Both tasks consist of novel and previously seen items, however they differ in the way the previously seen items were encoded. Items are consciously encoded when participants study them at an earlier phase in the experiment. Items are unconsciously encoded when "old" items are those which were previously seen in the priming task itself. For example, a priming task may consist of two trials where participants classify pictures as "man-made" or "not man-made". In the first trial all pictures seen by the participant are new. In the second trial, some of the pictures presented were previously seen in the first trial. Vuilleumier, Schwartz, Duhoux, Dolan, & Driver (2005) found that having participants explicitly memorize items produced a larger priming effect than if the items were presented in an earlier classification trial (unconsciously encoded).

For this study, one way priming was measured using a classification task with novel items and previously memorized items from the Old/New effect task (old items). Since the primary goal of this study was to differentiate individuals feigning cognitive deficit from honest responders (intact and brain injured), a task which maximizes a priming effect was preferable in order to help ensure that a group that may have a reduced priming effect such as a brain injured group would show a primed response to "old" items. In addition, within each trial there should be a "repetition effect" related to additional presentations of a picture. As mentioned earlier, repeated items produce larger ERP waves. The repetition priming task used in this study allowed for the comparison of the $1^{st}$ presentation, $2^{nd}$ presentation, and $3^{rd}$ presentation of an item.

*ERP Recording*

The electroencephalogram (EEG) was recorded from 32 scalp electrodes using an electrode cap with Ag/AgCl inserts. All scalp electrodes were referenced to an electrode placed between CZ and CPZ and later re-referred to linked mastoids for data analysis or referenced to the linked mastoids initially. In order to detect eye blink artifacts which would distort the ERP data measured at the scalp, horizontal electrooculogram (HEOG) and vertical electrooculogram (VEOG) were recorded with two pairs of electrodes, one electrode placed between the eyebrows and the other below the left eye, and another pair placed on the outer edge of each eye. EEG was recorded continuously during the tasks. EEG signals were filtered with a bandpass of .1~40 Hz and sampled at a rate of 500 Hz. Electrode impedances did not exceed 10 kilohms, ensuring a good connection between the electrode and scalp for an accurate ERP recording. Averaged ERPs were formed offline from correct-response trials, incorrect-responses trials, and all responses for the Old/New and the repetition priming tasks free of ocular and movement artifacts ($>\pm 75$ mV). For the Old/New task each scalp site resulted in two separate ERP waveforms (old and new). For the repetition priming task, each scalp site produces six separate ERPs ($1^{st}$ presentation, $2^{nd}$ presentation, and $3^{rd}$ presentation of new items [never seen before by participants] and $1^{st}$ presentation, $2^{nd}$ presentation and $3^{rd}$ presentation of old items [previously memorized in the Old/New task]). The mean numbers of individual trials per waveform that were free from movement or eye blink artifacts which distort the ERP recording were calculated. Generally, the average number of trials necessary for an accurate waveform for each condition is approximately 30-40.

*Procedure*

Participants completed demographic, BDI-II and WTAR questionnaires/tests

under standard instructions when they arrived at the lab. Participants with a history of brain injury also filled out a release of information for specific information about their brain injury from their medical records (date of injury, name of hospital, neuroimaging results for head (CT, MRI, or PET), duration of LOC, all Glasgow Coma Scores (GCS), neurosurgical interventions, history of previous brain injury, and discharge summary). A standard eye exam was administered to all participants in order to make sure their vision was 20/20 or corrected to 20/20. Participants in the non-brain injured and malingering groups were assigned to groups to match the age, education, and gender of participants in the TBI group as much as possible. Prior to beginning the TOMM-C, participants were given one of two sets of instructions, honest and brain injured groups received instructions to give their best effort on the three tasks and participants in the malingering group were instructed to pretend that they have memory problems as a result of an accident (see Appendix A for group instructions). After reading the instruction set, the RA asked each participant to paraphrase his/her understanding of how to respond to the tasks and did not move to the 1st task until the participant was clear about how to respond. Next, participants were asked to read over the instructions for the 1st task (TOMM-C) and there were two sample questions administered to the participant in front of the RA to ensure that the person knew the instructions for the task and the response instruction set (see Appendix B for instructions for the three tasks). Participants were administered Trial 1 and Trial 2 of the TOMM test on the computer in a sound-proof room. During the standard 15 minute delay between Trial 2 and the Retention Trial (T3) of the TOMM, two research assistants began to put a 32-electrode cap on each participant.

After completing the Retention Trial of the TOMM, the RAs finished putting on the EEG cap (approximately 10-15 minutes). After the cap was put on, the participants read the instructions for the 2nd task (Old/New task) and the RA answered any questions about the task. Participants then studied 100 new line drawings that were different from the TOMM pictures on a computer screen presented for 5 seconds each. The RA watched the participant take a brief practice trial (approximately three minutes) of the Old/New task to ensure that the participant understood the instructions for the 2nd task and the instructions regarding blinking for the EEG analysis. The participant then took the Old/New test. The third task was a classification task used to measure repetition priming where participants decided if the line drawing on the computer screen was "man-made" or "not man-made". The participant was asked to read the instructions and the RA again answered any questions about the 3rd task (Repetition Priming). There were three trials of this task. The computer response keys (i.e., right hand or left hand were "memorized" and "man-made" key) were counterbalanced for participants. In addition, the RAs were not blind to the condition for HON and MAL; it was not deemed necessary for this study where the three tasks to be compared were computer administered and not subject to any bias effects by the RA.

Chapter Three
Results
All behavioral (accuracy and RT) and ERP analyses were conducted on data from
16 participants in the HON group, 16 in the MAL group, and 15 in the TBI group (47
participants total). As shown in Table 1, based on ANOVAs or $\chi^2$ as appropriate, there
were no significant differences between groups (all $p$'s > .05) for any demographic
variables, including age, education, IQ estimate (WTAR), acute emotional state (BDI-II),
gender, race and marital status. The remaining results will be presented in four major
sections. First (1), the standard malingering variable of test accuracy from the three tasks
(TOMM-C, Old/New task, and Repetition Priming task) will be analyzed. Next (2), RT
will be analyzed from the same three tasks in order to determine if this newer variable has
validity for detecting MNCD. The third section (3) will evaluate the physiological
measure, ERP, as a malingering detection method in the Old/New and Repetition Priming
tasks. Then (4), the best variables from each previous section will be entered into a
stepwise logistic regression analysis in order to determine which are the most effective
predictors of malingering and if the newer variables of RT and ERP have incremental
classification power in predicting honest (clinical and non-clinical groups) vs.
malingering status. In the first four sections listed above, two comparisons will be made
related to clinical malingering (TBI vs. MAL) and lie detection in healthy adults (HON
vs. MAL). Both comparisons are relevant to the purpose of this study, mainly the
detection of MNCD. Finally, (5) the ERP results for the HON vs. TBI comparison will
be summarized separately. Although this comparison is not directly related to the
hypothesis of this study, the results have implications for research using ERPs to detect
MNCD.

*1. Accuracy Results*

To determine whether the three study groups obtained different accuracy rates, an
initial multivariate analysis of variance (MANOVA) was performed on the TOMM-C,
Old/New task, and Repetition Priming task. Results indicated a significant overall effect
(Wilks' 7 = .074, $F(16, 72) = 12.06$, $p < .001$). Unless otherwise mentioned, all post hoc
comparisons were conducted using Tukey's HSD test.

*1.1 TOMM-C*

TOMM accuracy results calculated as the percent correct out of 50 items in a trial
were analyzed using ANOVA to compare the three groups on Trial 2 and the Retention
Trial (Trial 3). The top section of Table 2 presents means and standard deviations for
TOMM accuracy data. There was a significant difference between groups for TOMM
Trial 2 accuracy rate, $F(2, 44) = 72.43$, $p < .0001$. There was also a significant difference
between groups for TOMM Trial 3 accuracy rate, $F(2, 44) = 77.43$, $p < .0001$. Post hoc
comparisons indicated that for both trials MAL had a significantly lower accuracy
(approximately 61 - 64%) than HON and TBI which both had almost perfect accuracy
(approximately 99 – 100%) ($p < .0001$ for both comparisons). These findings indicate
that the accuracy scores from this standard malingering test were successful in
documenting differences between MAL and both honest responding groups (HON and
TBI).

*1.2 Old/New Effect Task*

The middle section of Table 2 presents accuracy data from the Old/New effect
task. Percent correct was analyzed using ANOVA and followed up by Tukey's HSD
tests for post hoc comparisons. There was a significant difference in accuracy for old
items, $F(2, 44) = 15.58$, $p < .0001$. Post hoc analysis revealed that HON and TBI had

Table 1

*Demographic Information for the Three Study Groups*

| Variable | HON | Group<br>MAL | TBI |
|---|---|---|---|
| Age | | | |
| M | 36.2 | 32.7 | 40.5 |
| (SD) | (12.2) | (12.8) | (11.7) |
| Education | | | |
| M | 15.4 | 15.7 | 14.3 |
| (SD) | (2.3) | (2.5) | (1.9) |
| WTAR | | | |
| M | 110.3 | 111.8 | 105.5 |
| (SD) | (11.3) | (7.8) | (6.9) |
| BDI-II | | | |
| M | 5.1 | 2.5 | 6.9 |
| (SD) | (7.6) | (4.2) | (6.3) |
| Gender (% Female) | 56.3 | 62.5 | 46.7 |
| Race (% White) | 68.8 | 93.8 | 100.0 |
| Marital Status (% Single) | 50.0 | 56.3 | 53.3 |
| N | 16 | 16 | 15 |

*Note*. HON = Honest Responders; MAL = Malingering Responders; TBI = Traumatic Brain Injury group; WTAR = Wechsler Test of Adult Reading; BDI-II = Beck Depression Inventory-II. Age and education given in years. No variables were significantly different ($p$'s > .05).

Table 2
*ANOVA Results for Group Differences in Accuracy Data*

| Test | Variable | Group | | |
| | | HON | MAL | TBI |
|---|---|---|---|---|
| TOMM-C | | | | |
| | T2*** | 99.9$^a$ (.5) | 64.0$^b$ (16.7) | 99.6$^a$ (1.5) |
| | T3*** | 100$^a$ (0.0) | 61.1$^b$ (17.7) | 99.8$^a$ (.7) |
| Old/New task | | | | |
| | Old items*** | 93.5$^a$ (5.6) | 64.0$^b$ (17.6) | 81.4$^a$(18.5) |
| | New items** | 87.1$^a$ (9.8) | 67.2$^b$ (14.5) | 74.5$^b$(18.3) |
| | All items*** | 90.3$^a$ (5.8) | 65.6$^b$ (12.2) | 75.2$^b$(17.0) |
| | Diff. score* | 4.5$^{ab}$ (7.7) | -2.3$^a$ (14.8) | 8.7$^b$ (9.9) |
| Repetition Priming task^ | | | | |
| | 1$^{st}$** | 88.1$^a$ (8.0) | 75.3$^b$ (11.3) | 83.7$^a$ (6.9) |
| | 2$^{nd}$** | 91.5$^a$ (6.1) | 79.4$^b$ (11.4) | 87.2$^{ab}$(8.6) |
| | 3$^{rd}$** | 89.2$^a$ (12.0) | 79.4$^b$ (11.8) | 88.1$^{ab}$(7.7) |
| | New** | 87.6$^a$ (9.3) | 76.6$^b$ (11.5) | 85.9$^a$ (6.3) |
| | Old** | 91.6$^a$ (7.0) | 79.4$^b$ (11.8) | 86.7$^{ab}$(7.4) |
| | All items** | 89.6$^a$ (7.4) | 78.0$^b$ (11.2) | 86.3$^a$ (6.6) |

*Note.* HON = Honest Responders; MAL = Malingering Responders; TBI = Traumatic Brain Injury group. M = mean; SD = standard deviation; TOMM-C = Test of Memory Malingering-computerized; diff. score = old – new correct items from the Old/New task. $^{ab}$denotes groups significantly different from each other; same letter means groups are not different from each other; ^RT to wrong responses was not listed since 2 groups had N $\leq$ 2; *denotes significant main effect $p < .05$; **$p < .01$; ***$p < .001$.

higher accuracy rates (approximately 94% and 81%, respectively) compared to MAL (64%) ($p$'s < .0001 and .01). The difference score (old items correct − new items correct) was also significantly different for the groups, $F(2, 44) = 3.81$, $p < .05$. Post hoc analysis revealed a significant difference indicating TBI had a significantly larger difference score than MAL ($p < .05$). HON was not significantly different from MAL for this variable. These results suggest that accuracy for studied (old) items might be a new indicator for identifying MNCD, and to a lesser extent, the difference score for studied and new items may also index malingering.

Also shown in the middle section of Table 2, two other Old/New task variables showed significant group differences in accuracy; the accuracy rate for new items and the accuracy rate for all items (old and new items combined). However, because HON was significantly more accurate than both MAL and TBI here, these variables are unlikely to be useful indicators of malingering versus genuine TBI. Thus, no further between-groups analyses on these variables will be presented.

*1.3 Repetition Priming Task*

The bottom section of Table 2 presents accuracy data for the Repetition Priming task; these data were also analyzed using ANOVA (and Tukey's HSD for post hoc comparisons). There was a significant difference such that HON and TBI had higher accuracy rates compared to MAL on several accuracy variables for this task. The accuracy for the three groups for the 1st presentation of an item, old and new, (120 items) was significantly different, $F(2, 43) = 8.21$, $p = .001$. Post hoc analyses revealed significant differences ($p$'s < .01 and .04) such that the accuracy rates for HON (88%) and TBI (84%) were significantly higher than MAL (75%). Accuracy rates also differed for new items and all items combined. The accuracy of participant responses for new items (180 items, 60 never seen pictures presented 3 times) was significantly different between groups, $F(2, 43) = 6.19$, $p < .01$. Post hoc analyses revealed a significant difference ($p$'s < .01 and .03) such that HON (88%) and TBI (86%) had a significantly higher accuracy rate than MAL (77%). The accuracy of participant responses for all items (360 items, 120 pictures [half new and half old] presented 3 times) was significantly different between groups, $F(2, 44) = 7.31$, $p < .01$. Post hoc analyses reveal a significant difference ($p$'s < .01 and .04) indicating that HON (90%) and TBI (86%) had a significantly higher accuracy rate than MAL (78%). These findings suggest that the accuracy for the 1st presentation of an item, new items, and all items had results that suggest they might be useful variables to detect MNCD (accuracy for HON and TBI was higher than MAL).

As also shown in the bottom of Table 2, three other variables had significant group differences in accuracy across groups; the accuracy for the 2nd presentation of an item, old and new, (120 items) the accuracy for the 3rd presentation of an item, old and new, (120 items) and the accuracy of responses for old items (180 items, 60 previously seen pictures presented 3 times). These variables will not be analyzed further here as they only show group differences between MAL and HON. As previously noted, variables that only show differences between HON and MAL would not be the best method to detect MNCD. Variables that show differences between both HON and TBI compared to MAL would be ideal methods to detect malingering.

*1.4 Effect Sizes for Accuracy Data*

Ideally for a variable to be successful at detecting MNCD, effect size (ES) data should show large differences among the MAL and HON/TBI groups. Cohen's $d$ is a standard measure of the separation between two group means. It is calculated by

dividing the difference between two group means by the pooled standard deviation of the groups ($d = (M_1 - M_2) / \sqrt{([SD_1^2*(N_1-1)] + [SD_2^2*(N_2-1)]/ [N_1+N_2-2])}$). ES data are presented as an absolute value of the difference between two groups. Cohen (1988) attempted to define guidelines for comparing ES data. In general, a small ES is considered to be Cohen's $d = .2$, a moderate ES is considered to be Cohen's $d = .5$, and a large ES is considered to be Cohen's $d = .8$. Table 3 presents the ES data for previously presented accuracy variables. For TOMM-C accuracy variables the ES for the difference between HON and TBI compared to MAL was very large (Cohen's $d = 3.1$ and $3.0$ respectively) and further supported the strength of this traditional method of detecting MNCD.

To a lesser extent ES data also supported the accuracy for old items from the Old/New task as one that strongly differentiated between the TBI and MAL groups (Cohen's $d = 1.0$) as well as between HON and MAL (Cohen's $d = 2.3$). Most of the Repetition Priming variables had similar patterns of ES differences. The average ES for all significant accuracy variables indicated that the largest difference occurred between HON and MAL (Cohen's $d = 1.4$), the TBI and MAL groups had a slightly lower but still large ES (Cohen's $d = .9$), and the lowest ES occurred when comparing HON and TBI (Cohen's $d = .5$). Overall, the TOMM-C accuracy was the most successful variable in distinguishing between HON and TBI from MAL. To a lesser degree, the accuracy for old items in the Old/New task and one of the Repetition Priming accuracy variables might also have some promise as a way to detect MNCD.

*1.5 Classification Statistics for Accuracy Data*

The classification accuracy of diagnostic tests is traditionally expressed using specific terminology. Sensitivity and specificity are parameters describing a test's classification rate. In the case of malingering tests, sensitivity is the percentage of people who are malingering who are identified by the test as malingering. Specificity is the percentage of people who are not malingering who are identified by the test as not malingering. Sensitivity and specificity describe the accuracy of a test, given that the criterion status of an individual is known (as in this simulation study). Hit rate is the overall classification accuracy of a test such that the numbers of people who are correctly identified by the test as malingering or not malingering are divided by the total number of people in the sample. Two more clinically relevant statistics are Positive Predictive Power (PPP) and Negative Predictive Power (NPP). PPP is the percentage of positive test signs (test results indicating a person is malingering) that correctly identify individuals who are malingering (PPP = [base rate (BR)*Sensitivity] / [(BR*Sensitivity) + (False Positive Rate*(1-BR))]). NPP is the percentage of negative test signs (test results indicating a person is not malingering) that identify honest responders (NPP = Specificity*(1-BR)/[(Specificity*(1-BR)+False Negative Rate* BR]) . Unlike sensitivity and specificity, PPP and NPP are partly dependant on the base rate of malingering in a population. As mentioned earlier, the base rate for malingering after mild head injury was estimated to be almost 40% based on a survey of forensic neuropsychologists (Mittenberg et al., 2002). This base rate for MNCD is important, because as noted above, along with sensitivity and specificity values, base rates are used to determine PPP and NPP.

The base rate for malingering in this study was approximately 50% when comparing either non-clinical honest (HON) and MAL groups or TBI (clinical honest) and MAL groups. Table 4 presents selected accuracy variables and the cutting scores (using a 50% base rate of feigning) which obtained the highest overall hit rate for each

Table 3
*Effect Size Data in Cohen's d (Significant Results for Accuracy Variables)*

| | | Groups Compared: ES (95% CI) | | |
|---|---|---|---|---|
| Test | Variable | HON v MAL | TBI v MAL | HON v TBI |
| TOMM-C | | | | |
| | T2 | 3.0 (2.0 - 4.0) | 3.0 (1.9 -3.9) | .3 (-.4 - 1.0) |
| | T3 | 3.1 (2.0 - 4.1) | 3.0 (1.9 – 4.0) | .4 (-.3 - 1.1) |
| Mean ES for TOMM Accuracy Variables: | | 3.1 | 3.0 | .4 |
| Old/New task | | | | |
| | Old | 2.3 (1.3 - 3.1) | 1.0 (.2 - 1.7) | .9 (.1 - 1.6) |
| | New | 1.6 (.8 - 2.4) | .4 (.1 - 1.6) | .9 (.1 - 1.6) |
| | All | 2.6 (1.6 - 3.5) | .7 (.1 - 1.4) | 1.2 (.4 - 1.9) |
| Difference score (old – new items) | | .6 (.1 - 1.3) | .9 (.1 - 1.6) | .5 (-.3 - 1.2) |
| Mean ES for Old/New Accuracy Variables: | | 1.8 | .8 | .9 |
| Repetition Priming task | | | | |
| | 1st | 1.3 (.5 - 2.1) | .9 (.1 - 1.6) | .6 (.2 - 1.3) |
| | 2nd | 1.3 (.5 - 2.1) | .8 (.0 - 1.5) | .6 (.2 - 1.3) |
| | 3rd | .8 (.1 - 1.5) | .9 (.1 - 1.6) | .1 (-.6 - .8) |
| | New | 1.1 (.3 - 1.8) | 1.0 (.2 - 1.7) | .2 (-.5 - .9) |
| | Old | 1.3 (.5 - 2.0) | .7 (.0 - 1.5) | .7 (.1 - 1.4) |
| | All | 1.2 (.4 - 2.0) | .9 (.1 - 1.6) | .5 (.3 - 1.2) |
| Mean ES for RP Accuracy Variables: | | 1.4 | .9 | .5 |

*Note*. HON = Honest; MAL = Malingering; TBI = Traumatic Brain Injury; TOMM-C = Test of Memory Malingering-computerized; T3 = Trial 3; RP = Repetition Priming; ES are absolute values of mean differences between group.

variable in the two major comparisons, MAL vs. HON and MAL vs. TBI. The most successful accuracy variables were selected based on ANOVA results suggesting group differences between MAL and the two honest responding groups (HON and TBI) and/or ES data indicating large effects. In the top half of Table 4 when comparing MAL to HON (non-clinical honest) it can be seen that the highest hit rates were obtained when using the TOMM-C Trial 3 accuracy and the old accuracy for old items in the Old/New task (hit rates of 100% and 88%, respectively). In addition, the NPP and PPP were also 100% for the TOMM and 88% for the accuracy for items in the Old/New task. Turning to the bottom half of Table 4, when comparing MAL to the TBI (clinical honest) group, once again the highest hit rates were obtained when using the TOMM-C Trial 3 accuracy and the accuracy for old items in the Old/New task (hit rates of 100% and 74%, respectively). The TOMM-C trial 3 accuracy also obtained 100% NPP and PPP. The accuracy for old items in the Old/New task obtained a 72% NPP and 78% PPP. The TOMM-C Trial 3 accuracy and the accuracy for old items on the Old/New task were the most successful variables in classifying both groups overall. Another promising variable from the Repetition Priming task was the accuracy for the 1$^{st}$ repetition of an item. This variable was one of the most successful accuracy variables and had a low false positive rate (high specificity) when comparing TBI and MAL. When comparing these groups with a cutoff score of < 76% correct sensitivity was moderate at 53%, but specificity was excellent at 93% for a 73% hit rate. The TOMM-C had perfect classification rates, and to a lesser extent the accuracy for old items from the Old/New task were successful accuracy variables; however as noted below, these two variables will pose technical challenges when evaluating the possible incremental utility of other test variables.

*2. RT Results*

To determine whether the three study groups had different RTs, an initial multivariate analysis of variance (MANOVA) was performed on the TOMM-C, Old/New task, and Repetition Priming task. Results indicated a significant overall effect (Wilks' 7 = .004, $F(70, 20) = 4.03$, $p < .001$).

*2.1 TOMM-C*

RT is not a commonly utilized variable to detect MNCD, although it is available from most computerized malingering tests. The TOMM-C allows for RT measures to be examined along with accuracy data. Group differences in RT were measured using ANOVA and due to the unequal variances for RTs in this task, a Welch statistic was used and Tamhane's T2 post hoc correction. The top section of Table 5 presents RTs for Trial 2 and Trial 3 for the TOMM-C. RTs for correct and all responses in Trial 3 as well as an overall RT (all responses for Trial 2 and Trial 3) differed significantly across the three groups. For the two significant RT variables the only group differences in RT were between HON and MAL. HON had faster RTs compared to MAL, the RT for TBI was not significantly different from either group. The RT to all items in Trial 3 was also significantly different for the three groups, $F(2, 26.19) = 6.62$, $p < .01$. Post hoc comparisons indicated that HON was significantly faster than MAL by slightly less than approximately 216 ms ($p < .05$). The "overall" RT (all items combining both Trial 2 & 3) was significantly different for the groups, $F(2, 28.37) = 4.00$, $p < .05$. Post hoc comparison revealed that the overall RT for HON was significantly faster than the RT for MAL by approximately 165 ms ($p < .05$).

Both significant variables showed group differences between HON and MAL; however, the TOMM RT variable, Trial 3 correct items, might be the most successful at classifying groups because MAL had much slower RTs than HON and TBI had faster

Table 4

*Classification Statistics for Accuracy Variables (Base Rate for Simulation Malingering =
50%)*

| Groups Compared | Sensitivity | Specificity | Hit Rate | NPP | PPP |
|---|---|---|---|---|---|
| **Honest (non-clinical) vs. Malingering** | | | | | |
| TOMM-C: T3 (< 45/50 cutoff) | 100 | 100 | 100 | 100 | 100 |
| O/N task: Old items (< 86% correct) | 87.5 | 87.5 | 87.5 | 87.5 | 87.5 |
| O-N dif (< 0 = malingering) | 46.7 | 81.3 | 64.0 | 71.4 | 60.4 |
| Repetition Priming task | | | | | |
| RP task: New items (< 85% correct) | 80.0 | 75.0 | 77.5 | 79.0 | 76.2 |
| All items (< 87% correct) | 80.0 | 75.0 | 77.5 | 79.0 | 76.2 |
| 1$^{st}$ repetition (< 84% correct) | 80.0 | 75.0 | 77.5 | 79.0 | 76.2 |
| **TBI (clinical) vs. Malingering** | | | | | |
| TOMM-C: T3 (< 45/50 cutoff) | 100 | 100 | 100 | 100 | 100 |
| O/N task: Old item (< 72% correct) | 62.5 | 80.0 | 74.4 | 71.9 | 77.5 |
| O-N dif ($\leq$ 6 = malingering) | 62.5 | 73.3 | 67.9 | 66.2 | 70.1 |
| RP task: New item (< 85% correct) | 80.0 | 66.7 | 73.4 | 76.9 | 70.6 |
| All item (< 78% correct) | 53.3 | 86.7 | 70.0 | 65.0 | 80.0 |
| 1$^{st}$ repetition (< 76% correct) | 53.3 | 93.3 | 73.3 | 66.6 | 88.8 |

*Note*. HON = Honest Responders; MAL = Malingering Responders; TBI = Traumatic Brain Injury group;
TOMM-C = Test of Memory Malingering-computerized; T3 = Trial 3; RT = reaction time; O/N task =
Old/New task; O-N = old – new items; RP task = Repetition Priming task; ms = milliseconds; NPP =
Negative Predictive Power; PPP = Positive Predictive Power.

Table 5
*ANOVA Results for Group Differences in RT (in ms) Data*

| | | Group | | |
|---|---|---|---|---|
| | | HON | MAL | TBI |
| Test | Variable | | | |
| TOMM-C | | | | |
| T2 | RT-correct responses | 1214 (497) | 1371 (645) | 1443 (448) |
| | RT-all responses | 1220 (515) | 1552 (654) | 1440 (441) |
| T3 | RT-correct responses | 1084 (266) | 2314 (2286) | 1394 (418) |
| | RT-all responses* | $1084^a$ (266) | $1632^b$ (635) | $1395^{ab}$ (417) |
| Overall RT-T2 & T3 all responses* | | $1152^a$ (341) | $1592^b$ (585) | $1418^{ab}$ (356) |
| | | | | |
| Old/New task | | | | |
| Old | RT-correct responses | 732 (63) | 797 (91) | 844 (91) |
| | RT-incorrect responses | 825 (167) | 791 (268) | 880 (266) |
| | RT-all responses* | $732^{a1}$ (67) | $797^{ab}$ (93) | $855^{b1}$ (93) |
| New | RT-correct responses | 793 (78) | 823 (120) | 937 (93) |
| | RT-incorrect responses | 837 (173) | 781 (269) | 862 (154) |
| | RT-all responses* | $798^{a2}$ (80) | $811^a$ (119) | $921^{b2}$ (84) |
| RT difference in ms (correct) new - old* | | $66^a$ (59) | $14^{ab}$ (54) | $66^b$ (84) |
| | | | | |
| Repetition Priming task^ | | | | |
| 1st | RT (ms)-correct responses | 722 (76) | 769 (117) | 840 (110) |
| | RT (ms)-all responses | 725 (78) | 761 (141) | 840 (102) |
| 2nd | RT (ms)-correct responses | 651 (86) | 710 (100) | 753 (99) |
| | RT (ms)-all responses | 653 (89) | 713 (105) | 756 (100) |
| 3rd | RT-correct responses | 647 (74) | 681 (89) | 755 (102) |
| | RT-all responses | 651 (78) | 682 (95) | 753 (100) |
| New | RT-correct responses** | $672^a$ (78) | $727^{ab}$ (100) | $783^b$ (96) |
| | RT-all responses | 674 (83) | 723 (121) | 785 (94) |
| Old | RT-correct responses* | $675^a$ (74) | $714^{ab}$ (101) | $782^b$ (96) |
| | RT-all responses | 678 (76) | 714 (105) | 781 (104) |

*Note*. HON = Honest Responders; MAL = Malingering Responders; TBI = Traumatic Brain Injury group. M =mean; SD = standard deviation; RT = reaction time. [ab]denotes groups significantly different from each other; same letter means groups are not different from each other; ^RT to wrong responses was not listed since 2 groups had N ≤ 2; *denotes significant main effect $p \le .05$; **$p < .01$.

RTs than MAL (although not significant).  This trend of faster RTs for TBI compared to MAL will be important in a later section when determining which variable should be used as the "best" RT, specifically when comparing this variable with RTs from the Old/New and Repetition Priming tasks where TBI has the slowest RT of the three groups.

*2.2 Old/New Effect Task*

The Old/New task also produces RTs to old and new items.  The bottom half of Table 5 shows RTs to old and new items for all three possible responses (correct, incorrect, and all).   Results were analyzed using a 2 (Novelty) X 3 (Group) repeated measures mixed ANOVA and the Greenhouse-Geiser correction was used to adjust the degrees of freedom and thereby compensate for potential violations in sphericity in a repeated measures analysis (this correction was employed for all repeated measures ANOVAs).  There was a significant Novelty X Group interaction for the RT variable using all responses, $F(2, 44) = 3.20, p = .05$.  Post hoc analyses indicated between group differences such that TBI responded a little over 100ms slower to new items than HON and MAL ($p$'s < .001).  TBI also responded significantly slower to old items compared to HON ($p < .0001$).  This finding indicates that using only the RTs to new or old items would not help detect MNCD; however when looking at a pattern of RTs from the Old/New task there were some promising data.  There were within group differences such that both HON and TBI were 66 ms slower in responding to new items than old items ($p$'s < .0001 and .01).  MAL showed no significant difference in RTs to new and old items ($p = .32$).   Overall, TBI had slower RTs compared to HON and MAL.  TBI and HON had RTs that mirrored the "Old/New effect" such that the RT to new items was slower than to old items.  MAL did not show any RT difference between new and old items.

*2.3 Repetition Priming Task*

The second page of Table 5 presents RTs for each variable from the Repetition Priming task.  Two variables showed significant differences between the HON and MAL groups using 2 (Novelty) X 3 (Repetition) X 3 (Group) repeated measures mixed ANOVA and the Greenhouse-Geiser correction.  There was a significant difference between the groups when examining RT for correct responses that were not relevant to the detection of MNCD.  There was a significant interaction (novelty x group) such that the groups had significantly different RTs to new and old items, $F(2, 43) = 3.80, p < .05$.  Post hoc analyses indicated that HON responded faster to new and old items (approximately 674 ms for both) than TBI (approximately 783 ms for both) ($p$'s < .001).  RT results for the Repetition Priming task suggest that HON group responded faster than TBI to new and old items.  Again, similar to the results from the Old/New task, HON was faster than TBI suggesting that this type of time limited task does not detect MNCD.  There were no group differences related to the repetition effect.

*2.4 Effect Size Data for RT Data*

Table 6 presents ES data for the TOMM-C, Old/New task, and Repetition Priming task.  As mentioned earlier, to be an effective malingering detection strategy the variables should show large differences between MAL vs. TBI and MAL vs. HON.  The smallest difference should occur when comparing HON and TBI.  Although the comparison is not relevant to the hypothesis of this study, the largest ES data were obtained general when comparing HON vs. TBI (Cohen's *d*'s ranging from .8 – 1.5).  At the top of Table 6 the mean ES data for the three significant TOMM-C RT variables followed the same pattern as the earlier presented accuracy data and suggested that there were large differences

45

between HON compared to the MAL (Cohen's *d* = .9).  There were smaller but still moderate ES when comparing MAL and TBI (Cohen's *d* = .5).  When ES data were calculated for RTs to the Old/New task as shown in the middle section of Table 6, the mean differences between the groups indicated that a larger difference occurred when comparing MAL and TBI (Cohen's *d* = .6) compared to HON and MAL (Cohen's *d* = ).  Finally at the bottom of Table 6, the ES data for RTs to old items from the Repetition Priming task followed the same pattern as the Old/New task suggesting that there were larger ES when comparing MAL and TBI (Cohen's *d* = ) compared to HON and MAL (Cohen's *d* = ).  The ES data supported the use of the TOMM-C RT as showing slightly more differences between MAL vs. HON and MAL vs. TBI.  The TOMM-C gives test takers unlimited time to respond, in contrast to both the Old/New and Repetition Priming tasks which have short fixed times to respond.  The longer time to respond appears to lead to longer RTs in malingerers.

*2.5 Classification Statistics for RT Data*

In the same manner as detailed in the section regarding classification statistics for accuracy variables, the mean differences between the three groups for RT variables were compared in order to identity the best variables to classify HON,  MAL, and TBI.  Preliminary analyses showed two variables to be the best to classify malingering, the TOMM-C Trial 3 RT for correct items and the old-new RT difference score from the Old/New task.  The TOMM-C variable was chosen because of the mean difference between the groups was the largest and this difference was supported by larger ES differences between MAL and HON, as well as MAL and TBI.  The old-new RT difference score was chosen because of the within group differences in RT to old and new items for HON and TBI.  The RT variables are presented in Table 7 along with the cutting scores which obtained the highest overall hit rate.  When classifying HON and MAL, the TOMM-C Trial 3 RT for items answered correctly was the most successful (hit rate = 78%).  When classifying TBI and MAL, the TOMM-C Trial 3 RT for items answered correctly was also the most successful at differentiating between groups with a hit rate of 71%.  Thus, this relatively new RT variable (TOMM-C Trial 3) was the most successful at differentiating between both honest (non-clinical and clinical) and MAL groups.

*3. Event-Related Potential (ERP) Results*

As noted earlier, physiological measures are one of the most recent methods used to detect MNCD.  ERPs measure electrical activity from the scalp.  The ERP method has shown potential to reveal new information about malingering.  ERPs are associated with various cognitive processes, which can be measured by the mean amplitude or latency of specific ERP components. For instance, P300 (also known as P3 or Late Positive Component [LPC]) stands for a positive going ERP component about 300–700 ms after stimulus presentation. There are several variables that can be examined when using ERP measures (e.g., ERP mean amplitude, peak latency, comparing large groups of electrodes, etc). Describing every significant finding is beyond the scope of this study.  There were several variables that suggested consistent group differences between HON and TBI; however, since this comparison is not directly related to the hypothesis of this study (detecting MNCD) these results will be summarized at the end of this section.  The primary goal of this study was to investigate whether there are group differences in ERP responses between MAL compared to TBI, and to a lesser extent MAL and HON (as a measure of lie detection in healthy adults).

As mentioned earlier, the ideal number of "artifact free" trials to create an average

Table 6

*Effect Size Data in Cohen's d ( Significant Results for RT)*

| | | Groups Compared: ES (95% CI) | | |
|---|---|---|---|---|
| Test | Variable | MAL v HON | MAL v TBI | TBI v HON |
| TOMM-C | | | | |
| | T3 (crct) | .8 (.0 – 1.5) | .6 (-.2 – 1.3) | .9 (.1 – 1.6) |
| | T3 (all) | 1.1 (.4 - 1.8) | .4 (-.3 - 1.1) | .9 (.1 – 1.6) |
| | Overall (T2 & T3 all) | .9 (.2 – 1.6) | .4 (-.4 – 1.1) | .8 (.1 – 1.5) |
| Mean ES for task with no time limit: | | .9 | .5 | .9 |
| Old/New (O/N) task | | | | |
| | Old (all) | .8 (.1 – 1.5) | .6 (-1.3 - .1) | 1.5 (.7 – 2.3) |
| | New (all) | .1 (-.6 - .8) | 1.1 (-1.8 - ⁻.3) | 1.5 (.7 – 2.3) |
| Mean ES for O/N task: | | .5 | .9 | 1.5 |
| Repetition Priming (RP) task | | | | |
| | New (correct) | .1 (-.8 - .6) | .1 (-.8 - .6) | 0 (-.7 - .7) |
| | Old (correct) | .4 (-.3 - 1.1) | .7 (-1.4 - .1) | 1.3 (.5 – 2.0) |
| Mean ES for RP task: | | .3 | .4 | .7 |
| Mean ES for task with time limit: | | .4 | .6 | 1.1 |

*Note*. HON = Honest Responders; MAL = Malingering Responders; TBI = Traumatic Brain Injury group; TOMM-C = Test of Memory Malingering-computerized; T2 = Trial 2; T3 = Trial 3; ES presented as absolute value of difference between groups.

Table 7

*Classification Statistics for RT Variables (Base Rate for Simulation Malingering = 50%)*

| Variable | Groups Compared | Sensitivity | Specificity | Hit Rate | NPP | PPP |
|---|---|---|---|---|---|---|
| | HON (non-clinical) vs. MAL | | | | | |
| TOMM-C: T3 RT (crct) (> 1575ms) | | 62.5 | 93.8 | 78.2 | 71.4 | 91.0 |
| O/N task: O-N RT dif (< -50ms = mal) | | 80.0 | 62.5 | 71.3 | 75.8 | 68.1 |
| | TBI (clinical honest) vs. MAL | | | | | |
| TOMM-C: T3 RT (crct) (> 1575ms) | | 62.5 | 80.0 | 71.3 | 68.1 | 75.8 |
| O/N task: O-N RT dif (< -76ms = mal) | | 81.3 | 60.0 | 70.7 | 76.2 | 67.0 |

*Note*. HON = Honest Responders; MAL = Malingering Responders; TBI = Traumatic Brain Injury group; TOMM-C = Test of Memory Malingering-computerized; T3 = Trial 3; RT = reaction time; O/N task = Old/New task; O-N = old – new items; ms = milliseconds; crct = correct responses; NPP = Negative Predictive Power; PPP = Positive Predictive Power.

ERP wave for each participant is approximately 30-40.  For the most part, the number of trials was in an acceptable range for each group.  The mean number of correct and all items averaged for each HON participant ranged from 46 to 60 for the Old/New task (70 items for both conditions) and 38-49 for each of the six conditions of the Repetition Priming task (60 items per condition).  There were far fewer trials averaged for the wrong responses for HON ranging from 5 to 14.  While this number is well below the ideal number of trials for an average wave, it allowed for cautious comparison with MAL and TBI that had higher error rates.  Examining the wrong responses is warranted despite the reduced number of trials, because it may reveal a difference related to malingering that is obscured when comparing correct and all responses.  The mean number of correct and all items averaged for each TBI participant ranged from 30 to 51 for the Old/New effect and 26-44 for each of the 6 conditions of the Repetition Priming task.  There were fewer trials averaged for the wrong responses for TBI, ranging from 9 to 28.  The mean number of correct and all items averaged for each MAL participant ranged from 27 to 51.  There were fewer trials averaged for the wrong responses for the malingering group ranging from 9 to 21.

In some instances, participants did not have 30 trials to average.  Specifically, the number of trials acceptable for the "wrong" response average wave was less than ideal, so results that may show group differences were only emphasized if there was confirmatory evidence of a difference seen when analyzing other incorrect responses (showing an effect across wrong responses regardless of the task) or similar data from the analysis for correct and all responses (showing an effect across all responses).  In addition, two participants had slightly less than 30 trials for one type of item for "correct" responses (26-27); however, for the purposes of comparing three very different groups (honest responding, malingering, and TBI), all participants were analyzed and the variation in the number of trials averaged for each group, condition, and response type was noted above as the maximum and minimum mean values.

For the two tasks used in this study (Old/New and Repetition Priming), the mean ERP amplitude was examined at several time intervals (approximately 100-300 ms in length) to identify different brain responses associated with MNCD.  Several analyses were conducted to explore group differences in ERPs to detect malingering.  First, the ERP responses from six midline electrodes were compared in order to see the general activity produced by the task.  Since ERP is sensitive to temporal changes in milliseconds, the midline electrodes offer a good representation of an average ERPs evoked by the task.  The midline (mid-sagittal) electrodes are located over frontal, central, and posterior areas and were examined for an initial ERP analysis.  After looking at the results for the midline electrodes, expanded analyses using groups of electrodes follow when necessary.  Second, since the majority of group differences observed in the midline analyses occurred at frontal electrode sites (e.g., Fz), a comparison of activity from the frontal and the posterior electrodes (15 electrodes each) was conducted to see if there continued to be group differences that were related to larger electrode regions.  Third, a hemisphere analysis was conducted based on the findings of Tardiff et al. (2000) reporting a left frontal effect in malingering responders only.  In order to determine differences between the right and left hemisphere electrodes, the 12 electrodes on the right and left sides of the scalp (midline excluded) were compared to determine whether there was any hemisphere effect associated with malingering.  Figure 9 represents a diagram of the electrodes used in the three analyses mentioned above.  Finally, peak analyses were performed to determine if there were any group differences in the peak

amplitude and latencies for three typical ERP waves in the traditional ERP literature for the Old/New effect, i.e., the N200, P200, and P300. To reiterate, the results presented reflect group differences that distinguished MAL from HON and/or TBI.

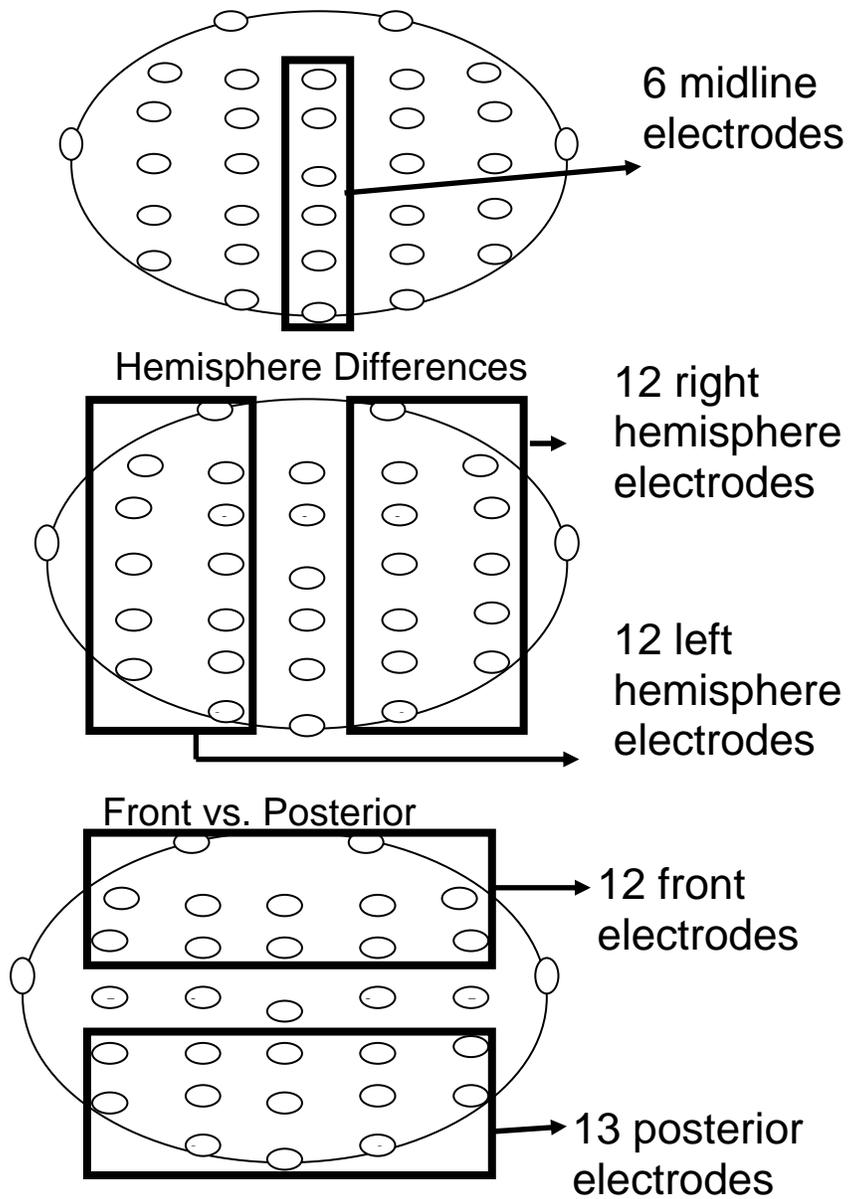*3.11 Old/New Task-ERP Mean Amplitude for Midline Electrodes*

The ERP responses to studied and new objects for the three groups were compared for the six midline electrodes in a 3 (group) X 2 (novelty) X 6 (electrode) mixed design repeated measures ANOVA for the three types of responses (correct, incorrect, and all responses). The Greenhouse-Geiser correction was used to identify significant interaction effects with the group variable and post hoc pairwise comparisons were conducted using Tukey's HSD test. In addition, the number of participants changed slightly (1-2 less) for some of the ERP analyses due to occasional "bad" channels, which led to a participant being dropped from an analysis. While this was a rare occurrence, a participant was not included in an analysis because a portion of the data was missing. The change is reflected in the degrees of freedom in the reported F-values. The mean ERP amplitude measured in mV was calculated for specific time intervals after stimulus presentation (0 ms). Visual inspection of ERP averaged waveforms was conducted for the three groups to determine time intervals that show the greatest difference among MAL, HON, and TBI. The time intervals examined for ERPs in the Old/New task were 0-100 ms, 100-250 ms, 250-400 ms, 400-650 ms, 650-750 ms and 750-900 ms. There was a significant interaction between Novelty X Group, $F(2, 44) = 3.3$, $p < .05$, across the midline electrodes for all responses at 400–650 ms interval, The "old/new effect" is a classic ERP finding reported in honest and normal healthy subjects, where ERP P300 responses to studied items are typically larger than those to new items. The following data are not presented in a table; however means and standard deviations are presented for each group and suggest that HON and TBI are more likely to show the old/new effect while MAL does not show any difference between old and new items in ERP response. For HON there was a significant difference in ERP mean amplitude such that studied tems ($M = 4.9$, $SD = 3.1$) evoked stronger ERPs than new items ($M = 2.9$, $SD = 2.2$) ($p < .01$). To a lesser extent, TBI followed the same typical trend of the old/new effect ($M_{old} = 3.2$, $SD = 2.7$; $M_{new} = 2.1$, $SD = 3.0$), even though the difference was not statistically significant ($p = .24$). MAL, however, had no difference in the late positive ERP responses between old ($M = 2.3$, $SD = 3.2$) and new items ($M = 2.2$, $SD = 2.7$), $p > .80$. Figures 10 & 11 show the ERP waves for the midline electrodes of the three groups for the Old/New task using all responses and the absent old/new effect in MAL is clearly visible. This finding suggests that the old/new effect is intact for HON, reduced but trending towards significant in TBI, and absent in MAL. Malingerers do not show the old/new effect indicating that feigning eliminates the expected ERP responses for new and studied items.

*3.12 Old/New Task: Old-New ERP Mean Amplitude Difference as Detection Strategy*

Based in part on the "old/new effect" results from the ERP mean amplitude for midline electrodes mentioned above, the difference between the ERP mean amplitude for old and new items at all 30 electrode sites across the scalp was calculated in order to determine if this pattern could be used to classify the groups as honest (TBI and non-TBI) or MAL at one electrode site. The classification of participants using individual ERP data was examined since visually inspecting ERP differences to determine if a person is malingering might be most useful for a clinician as a quick way to identify MNCD compared to using complicated statistical analyses, which requires additional time and

Figure 9

A diagram of electrode groupings for the three major ERP wave segment area analyses (midline, hemisphere, and frontal vs. posterior)



6 midline electrodes

Hemisphere Differences

12 right hemisphere electrodes

12 left hemisphere electrodes

Front vs. Posterior

12 front electrodes

13 posterior electrodes

effort. Initially, the electrodes that indicated the largest group differences were selected, and a card with the old and new ERP waves were made for each participant.

The six electrodes that appeared to show group differences in the old/new effect were FP2, Fz, F7, CP3, TP7, TP8 (3 frontal and 3 coronal midline electrodes). After pilot data were gathered from each electrode, the electrode Fz had a much higher hit-rate for classification than the other 5 electrodes. A person blinded to the purpose of the test was asked to classify honest responders by visually identifying the old/new difference in the interval of 300-700 ms (a typical P3 ERP component). When using this method with all three groups (TBI and HON initially collapsed into one "honest" group), the sensitivity was 67% and the specificity was 68%. These findings are a somewhat low rate of detection, so the two honest responding groups (HON and TBI) were compared to MAL separately in order to see if ERPs were more effective with one type of honest group in detecting MNCD.

The effectiveness of this method to detect malingering was more successful when comparing the MAL and TBI groups. The greatest difference occurred in the specificity of the test. When the "not malingering" group was TBI the specificity was 73%, but it was only 63% for HON. This suggests that individual ERP waves from the Old/New task was slightly more successful in distinguishing malingerers from TBI than in identifying malingerers from a non-brain injured honest group. However, using the visual inspection of ERPs from the Old/New task to detect malingering yields disappointing accuracy (hit rates of approximately 65% and 70%). There is a great deal of variation in ERPs to old and new items from individual participants which reduces the ability to use visual inspection of ERPs to quickly classify individual as malingering. Using larger groups of electrodes to find within group variations in ERP mean amplitude may be a better way to detect MNCD using ERPs.

*3.13 Old/New Task-Frontal vs. Posterior Electrodes Analysis*

The differences between ERPs for frontal vs. posterior electrodes were also examined (using the same time intervals mentioned above for the midline analysis). In the frontal electrodes HON had larger mean ERP amplitude than MAL, and MAL had larger mean amplitude than TBI (i.e., HON > MAL > TBI). This suggests that in frontal electrodes malingering reduces ERP mean amplitude but not as low as in participants who have had a brain injury. In addition, both HON and TBI show significant differences when examining frontal vs. posterior electrodes. HON has larger mean amplitude for frontal electrodes; TBI has larger mean amplitude for posterior electrodes. Unlike HON and TBI, MAL has no difference in activity between frontal and posterior electrodes. Malingering appears to have reduced variation across the scalp.

The number of participants included in these analyses changed to 15 for the HON and MAL groups and 14 for TBI due to exclusion from the analysis when at least one of the electrodes in the frontal or posterior groupings was not recording properly for a particular participant. A significant group difference occurred in the frontal electrodes. For all three responses (correct, incorrect, and all), MAL had larger mean amplitudes than TBI and a moderate ES difference (mean Cohen's $d = .4$). In addition, HON had larger ERP mean amplitudes than MAL and a large ES difference (mean Cohen's $d = .9$). When healthy individuals malingered the frontal electrodes show reduced ERP mean amplitude. Malingering appears to diminish ERP mean amplitude in frontal electrodes, but this type of response is still larger than the ERPs for individuals with brain injury.

The most interesting finding, when comparing the mean ERP amplitude of frontal vs. posterior electrodes, was the within group differences that occurred during the

Figure 10
ERP waves for the Old/New task for honest, malingering and TBI groups - frontal
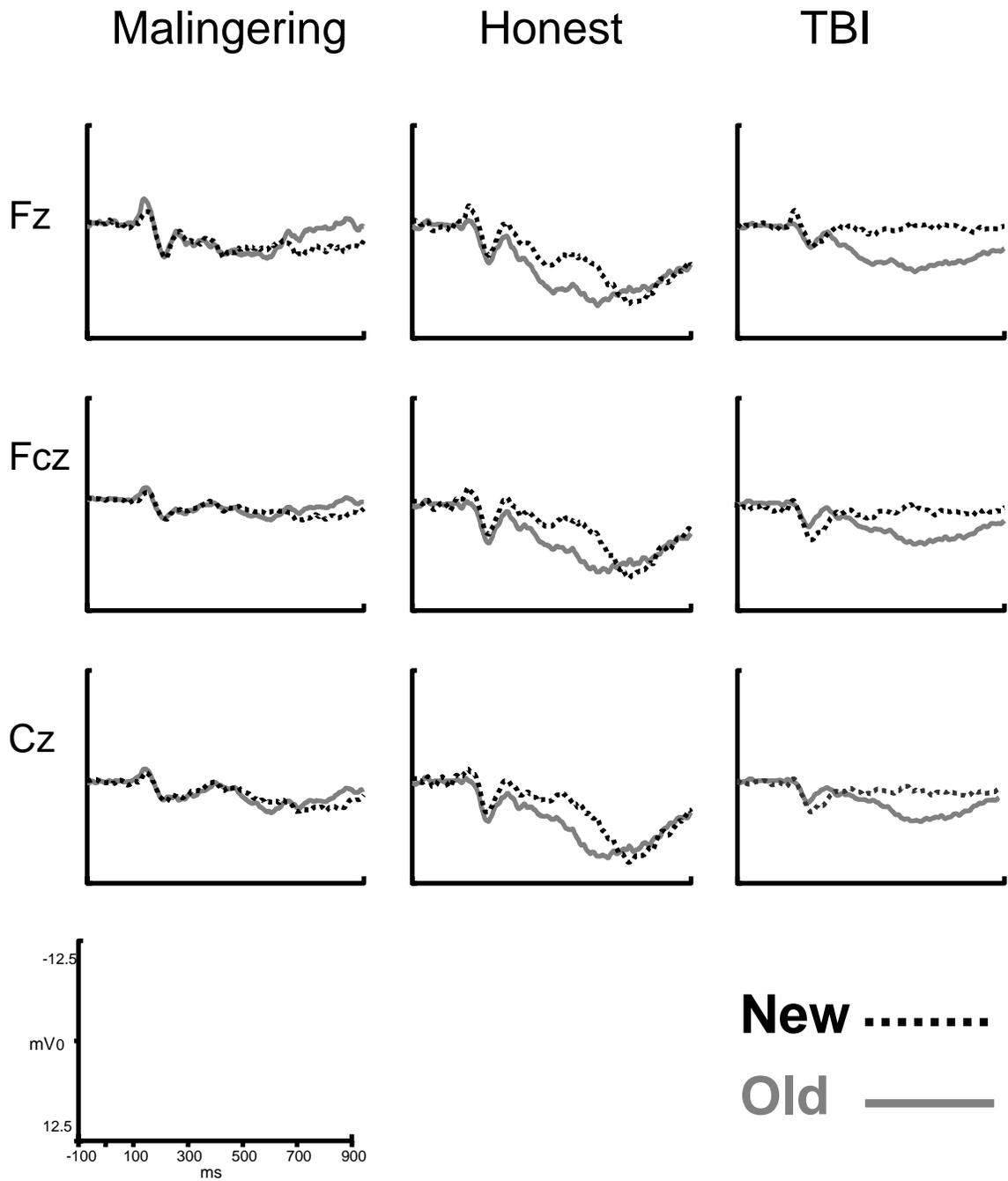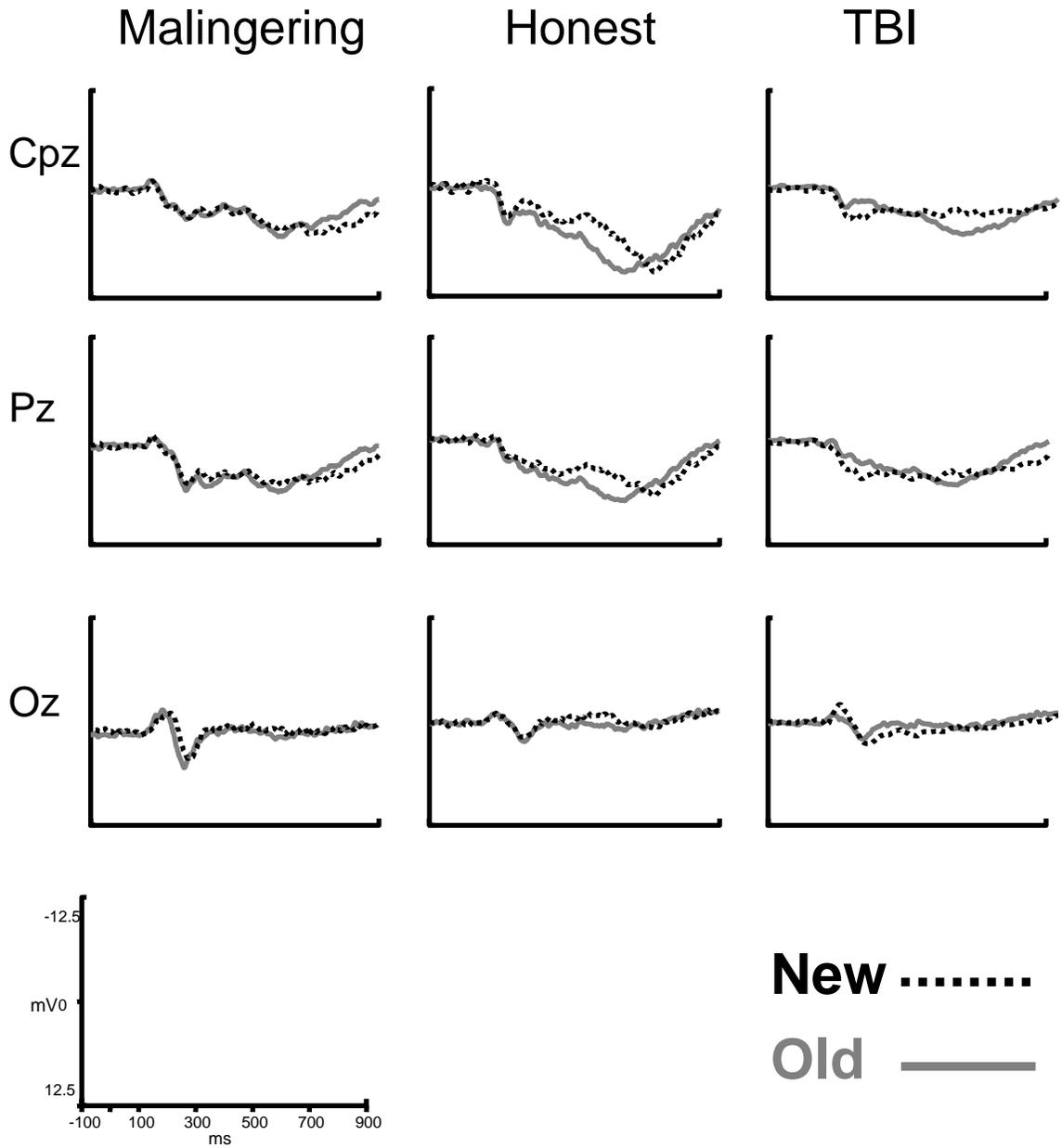midline electrodes (all responses)

Figure 11
ERP waves for the Old/New task for honest, malingering and TBI groups - posterior midline electrodes (all responses)

intervals of 400–900 ms.  Figure 12 represents the frontal vs. posterior electrode differences in recorded electrical activity among the groups using a topographical map across the scalp during the interval 400-650 ms.  The topographical map shows that when comparing activity from frontal and posterior electrodes within each group, HON had the strongest activity in the frontal area, TBI had the strongest responses in the posterior area, and MAL showed no significant difference between frontal and posterior activity.  Table 8 presents the mean ERP amplitude for each group.  The results indicated MAL showed no difference between frontal and posterior electrodes ($M_{frontal}$ = 2.4; $M_{posterior}$ = 2.2).  HON had larger mean ERP amplitude for frontal electrodes ($M_{frontal}$ = 4.7) compared to posterior electrodes ($M_{posterior}$ = 2.5).  Finally, individuals with TBI responding honestly use more posterior resources ($M_{posterior}$ = 3.0) compared to frontal ($M_{frontal}$ = 1.0).   The pattern of activity (frontal vs. posterior) for the Old/New task is important to note, because all three groups look different and the difference might help identify malingerers.  These findings suggest that non-brain injured individuals responding honestly show more frontal activity compared to posterior; however when a person is malingering the frontal responses are reduced (less positive ERP amplitude).  In addition, people with TBI have greater activity in posterior electrodes compared to frontal.  Using a difference score may help identify malingering, because in malingerers would be expected to have the smallest absolute value (close to "0").

     *3.14 Old/New Task-ERP Peak Amplitude and Latency*

     Peak latency and amplitude were examined for three major ERP components for the Old/New task.  Data were obtained for the N200 using the interval 100-300 ms, for the P200 the interval was 150-350 ms, and for the P300 or LPC an interval of 300-800 ms.  Group differences in these classic components may reveal how malingering distorts normal cognitive markers in a task, which can be evidence of related cognitive changes occurring when a person malingers that could be used as a marker of feigning.

     The most relevant result was obtained for the P200 peak amplitude using all responses.  There was a significant interaction for Novelty X Electrode X Group, $F(10, 210)$ = 2.69, $p < .05$.  Table 9 presents P200 peak amplitude data showing group differences in amplitude for new vs. old items.  MAL again showed no differences between new and old items, while both HON and TBI had larger ERP amplitude for old items compared to new.  At electrodes Fz-Pz (the front and central midline electrodes), HON had significantly larger peak amplitude for old items (average peak amplitude for the 5 electrodes = 7.16 mV; range 6.63 – 7.86) compared to new items (average amplitude for the 5 electrodes = 5.99 mV; range 5.6 – 6.44) ($p$'s < .01 - .05).  TBI also had a significantly larger peak amplitude for old items ($M = 5.82$, $SD = 3.27$) compared to new items ($M = 5.05$, $SD = 2.95$) for electrode Oz ($p < .05$).  MAL had no significant differences in P200 peak amplitudes for old and new items ($p$'s > .05).  Thus, for the P200 peak amplitude at the frontal and central electrodes there were larger peak amplitudes for old items for HON.  This P200 pattern of larger peak amplitude for the old items than for the new was also true for TBI at the most posterior midline electrode, Oz.  This finding suggests that while HON and, to a lesser extent, TBI processed the old and new items differently early on (150 - 350 ms after a picture was presented), MAL did not show the classic old/new effect in P200 peak amplitude.  Once again, the old/new effect appears to have been masked by the malingering response.

     *3.15 Old/New Task-Right vs. Left Hemisphere Analysis*

     As mentioned above, based on prior research reporting a left hemisphere effect for malingering responders using ERP, the ERPs for the right and left hemispheres were

Table 8

*ERP Mean Amplitude (mV) Frontal vs. Posterior Electrode Analyses from Old/New Task*

| Interval | | | HON | Group<br>MAL | TBI |
|---|---|---|---|---|---|
| | Overall mean for frontal electrodes | | 4.7 | 2.4 | 1.0 |
| | Overall mean for posterior electrodes | | 2.5 | 2.2 | 3.0 |
| 250-400ms | Front/Post X Group (all) | front | 3.4 (2.9) | 2.2 (2.8) | $1.5^1$ (3.5) |
| | | post | 2.6 (2.1) | 2.6 (2.0) | $3.9^2$ (2.2) |
| | Front/Post X Group (correct) | front | 3.2 (3.0) | 1.9 (2.9) | $1.5^1$ (3.5) |
| | | post | 2.5 (2.1) | 2.5 (2.2) | $3.6^2$ (2.1) |
| | Front/Post X Group (wrong) | front | 4.0 (3.3) | 2.7 (3.1) | $1.5^1$ (3.9) |
| | | post | 3.0 (2.4) | 2.7 (2.0) | $4.1^2$ (2.5) |
| 400-650ms | Front/Post X Group (all) | front | $4.7^{a1}$ (2.5) | $2.0^{ab}$(3.2) | $1.0^{b1}$(3.6) |
| | | post | $2.9^2$ (2.2) | 2.0 (2.1) | $3.6^2$ (2.2) |
| | Front/Post X Group (correct) | front | $4.8^{a1}$ (2.6) | $1.8^b$ (3.5) | $1.5^{b1}$(3.7) |
| | | post | $3.1^2$ (2.2) | 2.3 (2.3) | $3.6^2$ (2.4) |
| | Front/Post X Group (wrong) | front | $3.3^1$ (3.6) | 2.2 (3.2) | $.6^1$ (3.7) |
| | | post | $.9^2$ (1.8) | 2.3 (1.9) | $3.4^2$ (2.0) |
| 650-750ms | Front/Post X Group (all) | front | $6.2^{a1}$ (3.6) | $2.7^b$ (2.9) | $.6^{b1}$(3.3) |
| | | post | $3.4^2$ (2.1) | 2.2 (2.1) | $2.7^2$ (2.4) |
| | Front/Post X Group (correct) | front | $6.4^{a1}$ (3.5) | $2.6^b$ (3.5) | $1.3^b$ (3.4) |
| | | post | $3.9^2$ (2.1) | 2.3 (2.3) | 3.3 (2.4) |
| | Front/Post X Group (wrong) | front | $5.3^{a1}$ (3.8) | $2.5^{ab}$(3.0) | $-.1^b$ (3.5) |
| | | post | $.8^2$ (2.3) | 1.7 (2.2) | 1.2 (2.5) |
| 750-900ms | Front/Post X Group (all) | front | $5.2^{a1}$ (3.4) | 2.8 (2.7) | $1.3^b$ (3.5) |
| | | post | $1.5^2$ (1.7) | 1.3 (2.0) | 1.9 (2.7) |
| | Front/Post X Group (correct) | front | $5.1^{a1}$ (3.6) | $2.3^{ab}$(3.2) | $1.6^b$ (3.9) |
| | | post | $1.9^2$ (1.8) | 1.3 (2.2) | 2.5 (2.8) |
| | Front/Post X Group (wrong) | front | $4.9^{a1}$ (4.0) | $3.0^{ab1}$ (2.8) | $.8^b$ (3.7) |
| | | post | $-.1^2$ (2.4) | $1.2^2$ (2.2) | 1.0 (2.3) |

*Note.* HON = Honest Responders; MAL = Malingering Responders; TBI = Traumatic Brain Injury group. M =mean; SD = standard deviation; mV = millivolts. [ab]denotes groups significantly different from each other; same letter means groups are not different from each other; [12]denotes within group variable significantly different; same number means variables not different; all *p*'s < .05.

Figure 12

Scalp topography for the Old/New task (400-650ms) using all responses for honest, malingering and TBI groups
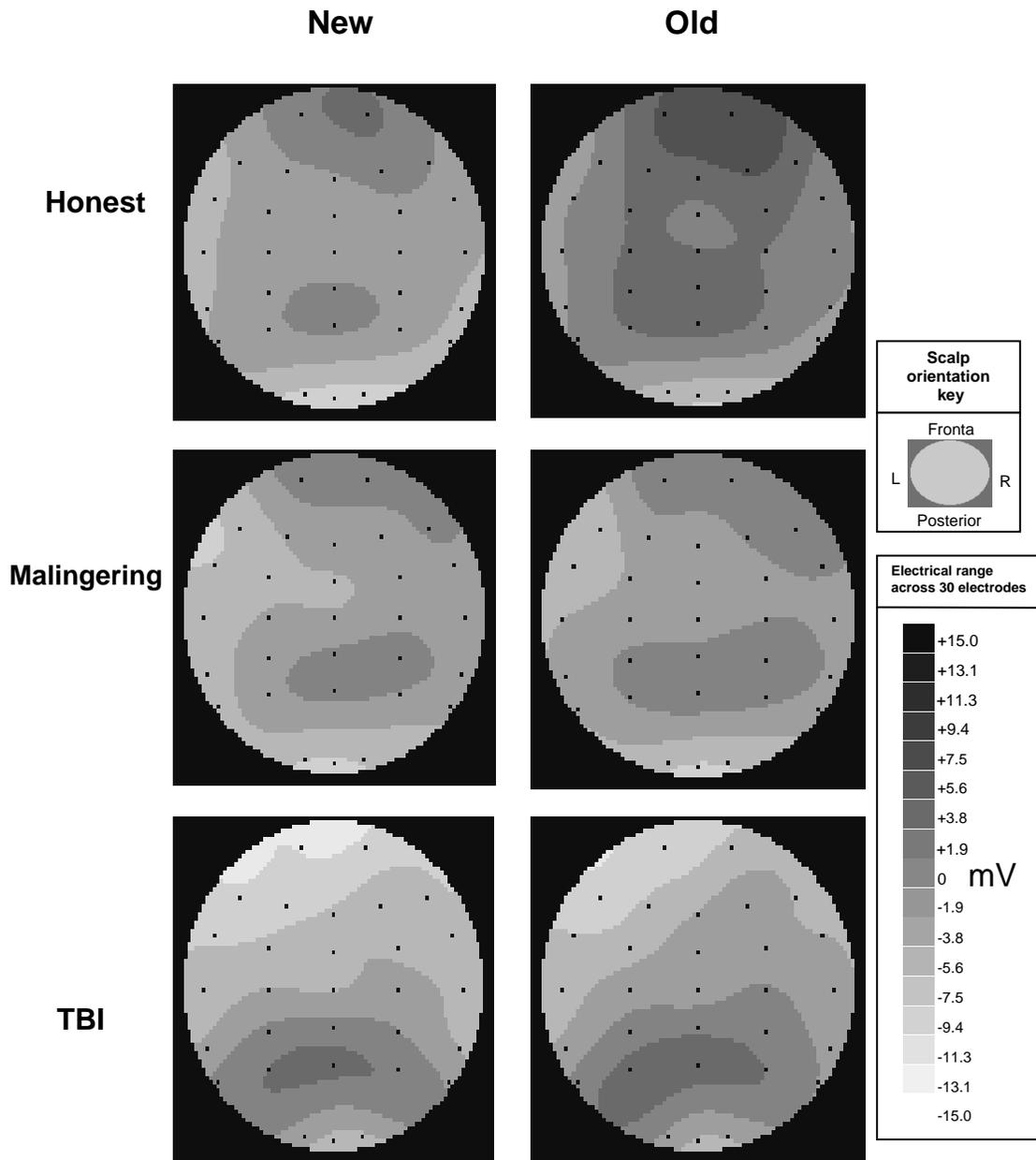
Table 9
*Old/New Task - P200 Peak Amplitude (mV) Analysis*

| | | | | Group | |
| | | | HON | MAL | TBI |
| --- | --- | --- | --- | --- | --- |
| significant interaction | | | | | |
| Novelty X Electrode X Group (all) | | | | | |
| | Fz | old | $6.9^1$ (3.2) | 6.1 (3.7) | 6.1 (3.1) |
| | | new | $5.7^2$ (3.0) | 5.5 (3.4) | 6.6 (3.1) |
| | Fcz | old | $6.6^1$ (3.4) | 6.3 (3.2) | 6.5 (3.7) |
| | | new | $5.6^2$ (3.1) | 5.5 (2.8) | 7.1 (3.0) |
| | Cz | old | $6.8^1$ (3.3) | 5.9 (3.0) | 6.6 (3.3) |
| | | new | $5.9^2$ (2.8) | 5.4 (2.3) | 7.1 (3.2) |
| | Cpz | old | $7.6^1$ (3.1) | 6.0 (3.3) | 7.2 (3.2) |
| | | new | $6.3^2$ (2.6) | 5.9 (2.8) | 7.5 (3.2) |
| | Pz | old | $7.9^1$ (3.0) | 7.3 (3.6) | 8.4 (3.0) |
| | | new | $6.4^2$ (2.9) | 6.5 (3.5) | 8.1 (3.4) |
| | Oz | old | 4.1 (3.1) | 4.3 (3.8) | $5.8^1$ (3.3) |
| | | new | 3.9 (3.2) | 4.3 (3.2) | $5.1^2$ (3.0) |

*Note*. HON = Honest Responders; MAL = Malingering Responders; TBI = Traumatic Brain Injury group. M =mean; SD = standard deviation; mV = millivolts. [ab]denotes groups significantly different from each other; same letter means groups are not different from each other; [12]denotes within group variable significantly different; same number means variables not different; all $p$'s < .05.

compared for group differences.  No significant differences related to hemisphere effects were found.  In order to specifically address the prior results from Tardiff et al. (2000), the two hemispheres were then divided into quadrants to attempt to identify more specific group differences (i.e., frontal-right, frontal-left, posterior-right, and posterior-left).  There were no significant differences related to malingering that were not accounted for in the frontal vs. posterior analysis.

*3.21 Repetition Priming Task-ERP Mean Amplitude for Midline Electrodes*

In general, the same ERP analyses were performed on the Repetition Priming task as were conducted on the Old/New task.  Mean ERP amplitude was examined at smaller intervals in order to determine where groups differed in response to repeated items along the midline of the scalp.  Similarly to the old-new task, comparisons of group differences in recorded electrical activity in the frontal vs. posterior electrodes were conducted.  Third, a peak analysis was performed to determine the peak amplitude and latency for the P300 reported in the ERP literature for the "repetition effect" (Rugg and Nagy, 1989; Paller, 2001).  The results from the peak analysis showed differences only related to the presence of brain injury (TBI compared to HON and MAL) so it will not be presented in detail here since the current focus is on detecting malingering.  In addition, a hemisphere analysis was not conducted for the Repetition Priming task due to the lack of a hemisphere effect associated with malingering in the Old/New task.  The ERP mean amplitude for the three groups were compared in a 3 (group) X 2 (novelty) X 3 (repetition) X 6 (electrode) mixed design repeated measures ANOVA for the three possible responses (correct, incorrect, all).  The intervals examined for the Repetition Priming task were 0-200 ms, 200-400 ms, 400-600 ms, 400-600 ms, and 600-900 ms.  There was no significant group X repetition effects for the midline electrodes.  Since the group differences were generally similar to those for Old/New task (i.e., mean ERP amplitude for HON > MAL > TBI) there is no table presenting these results.

*3.22 Repetition Priming Task-Frontal vs. Posterior Electrode Analysis*

The ERP differences in frontal vs. posterior electrodes were examined for the Repetition Priming task, using the original intervals from the midline area analyses. Table 10 presents group differences in the "repetition effect". When examining all responses during the interval of 400–600 ms, a significant interaction was found, i.e., Novelty X Repetition X Group, $F(4, 84) = 2.78$, $p < .05$.  Post hoc analyses revealed MAL only showed significant repetition effect for studied objects ($p = .05$).  TBI only showed a significant repetition effect for new items TBI ($p < .0001$).  These findings are different than the expected repetition effect for both new and old items as seen in HON ($p$'s $< .05$).  Once again, all three groups have different ERPs.  These results suggest that honestly responding, non-brain injured adults show the typical "repetition effect" for both old and new items.  Individuals with brain injury only show the repetition effect for new items indicating that the presence of a brain injury impairs normal processing of old items.  Curiously, MAL only shows the "repetition effect" for studied items, indicating that malingering may effect new items more than studied items during an unrelated (classification) task (see Figure 13).  Larger ERPs occur to repetitions due to increased synchrony in neurons that fire in response to the repeated item.  The lack of repetition effect in MAL for new items may be the result of malingering eliminating the repetition effect when there was not a prior exposure to an item (previously studied in an earlier task).  Studied items may have established an efficient neural network that results in a repetition effect despite the malingering response.  In contrast, both HON and TBI show the classic repetition effect for new items.

59

Figure 13

ERP waves from the Repetition Priming task for honest, malingering and TBI groups at electrode Cpz (all responses).
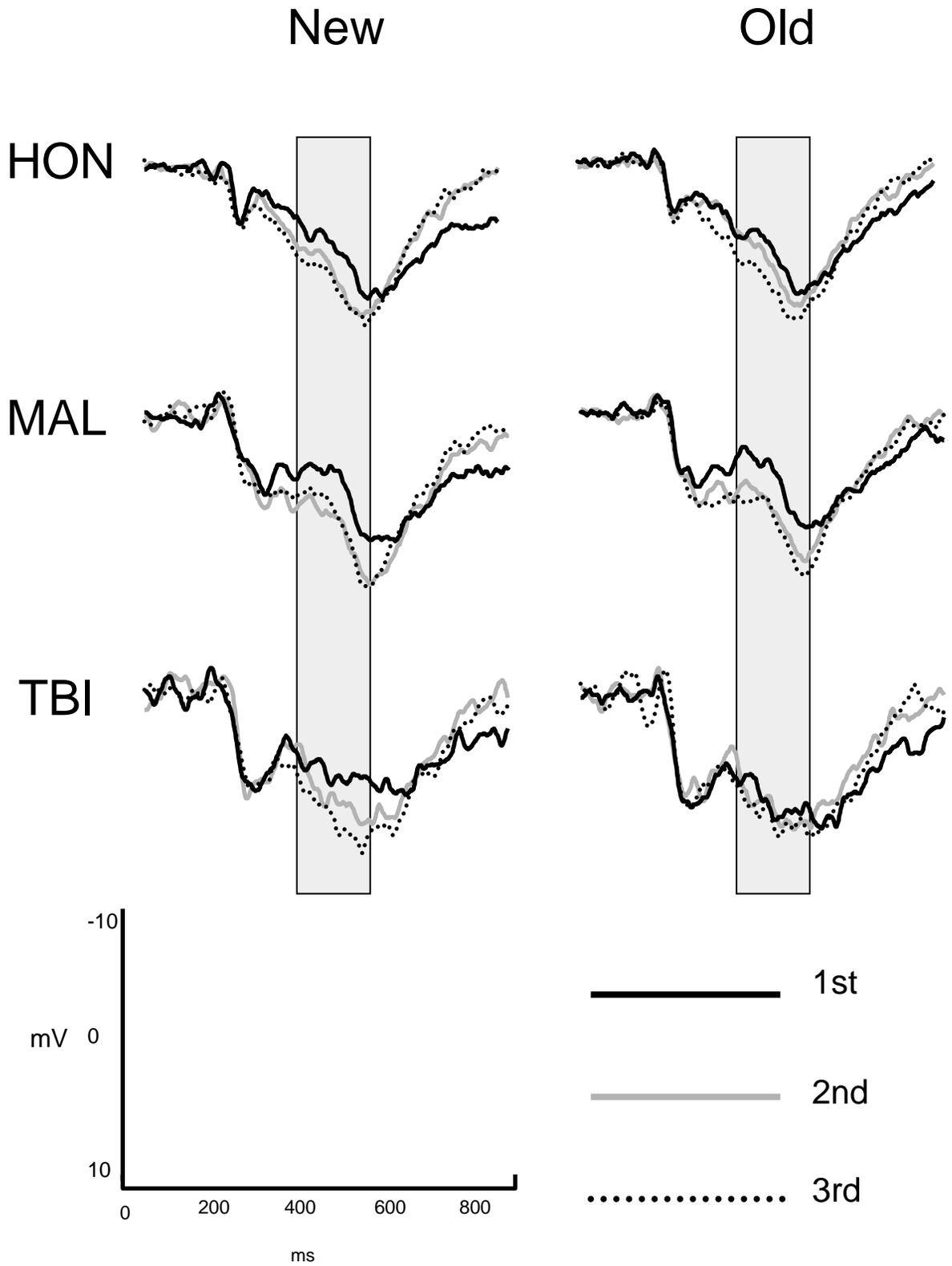
*3.23 Effect Size Data for ERP Results*

ES data was calculated for the Old/New and Repetition Priming tasks as an indication of the strength of the differences for significant ERP variables as indicated by the ANOVA results.  There were two comparisons that were relevant to the detection of malingering, BI vs. MAL and HON vs. MAL.  The ES data show different patterns of results depending on the group being compared to MAL.

*TBI vs. MAL-Clinical Comparison to Detect Malingering.*
Table 11 presents ES data comparing TBI vs. MAL and generally indicated small differences between MAL and TBI in mean ERP amplitudes for the Old/New task. When analyzing the mean ERP amplitude for the midline electrodes, the magnitude of the difference between TBI and MAL was small (mean Cohen's $d$ = .3).  When examining the difference in ERP mean amplitude for the frontal and posterior electrodes, the size of the difference between MAL and TBI was small to moderate (Cohen's $d$ = .4). Therefore, in general the mean ERP amplitudes for MAL was somewhat larger than TBI and the difference produced a small to moderate effect size for the Old/New task.  The largest ES was seen in the mean ERP amplitude for frontal electrodes during the interval of 650-750 ms for wrong responses (Cohen's $d$ = .8).  Table 12 presents ES data from the Repetition Priming task and also revealed small to moderate standardized group differences.  There was a small difference between TBI and MAL (Cohen's $d$ = .3) at midline electrodes.  The differences in ERP mean amplitude comparing frontal vs. posterior activity also show that there was a moderate difference between MAL and TBI (Cohen's $d$ = .5).  Generally, there were small to moderate differences in mean ERP amplitude when comparing TBI and MAL.  Comparing group means, individuals feigning memory deficits have slightly larger mean ERP amplitude compared to those with true brain impairment.  The Repetition Priming variable that had the largest ES when comparing MAL and TBI was the mean ERP amplitude for the 2$^{nd}$ presentation of an item during the interval of 0-200 ms using wrong responses (Cohen's $d$ = 1.0).  The group means for this variable suggest that MAL had negative-going waves for this variable while TBI had more positive-going waves in this early component.

*HON vs. MAL-Lie Detection in Healthy Adults.*
The mean ERP amplitude comparing the HON vs. MAL from the Old/New task indicated that HON had larger ERPs than MAL and the strength of the difference was large (large ES).  As mentioned above, Table 11 presents the ES data for the Old/New task.  When analyzing the midline electrodes, HON had a larger difference in ERP mean amplitude compared to MAL (Cohen's $d$ = 1.0).  The difference in ERP mean amplitude for the frontal and posterior electrodes indicated another large, overall difference between HON and MAL (Cohen's $d$ = .9).  Therefore, in general the difference between HON ERP responses and MAL was strong in the Old/New task.  Table 12 presents the ES data from the Repetition Priming task revealed a similar pattern of group differences found in the Old/New task.  The overall strength of the difference in ERP mean amplitude from the Repetition Priming task suggested that there was a large difference between HON and MAL (Cohen's $d$ = .8).  The differences in ERP mean amplitude comparing frontal vs. posterior activity also show a large difference in ERP mean responses between HON and MAL (Cohen's $d$ = .9).  Parallel to the results from the Old/New task, group means indicated HON had larger ERP mean amplitude than MAL.  Overall, as a measure of lie detection in healthy adults, ERPs show consistent and clear differences where malingerers show lower ERP mean amplitude regardless of the type of memory task.

Table 10

*Repetition Priming Task Group Differences in Repetition Effect - Mean Amplitude Analysis (mV)*

| Interval | significant interaction | | Group HON | MAL | TBI |
|---|---|---|---|---|---|
| 400-600ms | Novelty X Repetition X Group (all) | | | | |
| | Old | 1st | $4.3^1$ (1.5) | $2.7^1$ (2.1) | 2.7 (2.1) |
| | | 2nd | $5.1^{a12}$ (2.6) | $3.7^{ab12}$ (2.7) | $2.5^b$ (2.8) |
| | | 3rd | $5.8^{a2}$ (2.6) | $3.7^{ab2}$ (2.6) | $2.5^b$ (2.6) |
| | New | 1st | $4.9^1$ (2.1) | 3.1 (2.8) | $1.9^1$ (2.6) |
| | | 2nd | $5.5^{a12}$ (2.5) | $4.1^{ab}$ (2.6) | $2.8^{b12}$ (3.0) |
| | | 3rd | $5.9^{a2}$ (1.9) | $3.6^b$ (2.7) | $3.8^{ab2}$ (2.7) |

*Note*. HON = Honest Responders; MAL = Malingering Responders; TBI = Traumatic Brain Injury group; M =mean; SD = standard deviation; mV = millivolts. [ab]denotes groups significantly different from each other; same letter means groups are not different from each other; [12]denotes within group variable significantly different; same number means variables not different; all $p$'s < .05.

Table 11

*Effect Size Data in Cohen's d (for Significant Results Midline Electrodes Old/New Task)*

| Electrodes Analyzed | | Groups Compared: ES (95% CI) | | |
|---|---|---|---|---|
| Variable | Interval | HON v MAL | MAL v TBI | HON v TBI |

**Midline Electrodes**

| | 400-650ms | | | |
|---|---|---|---|---|
| Old – New difference score (all) | | 1.2 (.4 - 1.9) | .7 (- .1 - 1.4) | .4 (-.3 - 1.2) |
| | 650-750ms | | | |
| Electrode X Group (all) | Fz | 1.2 (.4 - 1.9) | .4 (- .3 - 1.1) | 1.5 (.7 - 2.3) |
| | Fcz | 1.0 (.2 - 1.7) | .6 (- .1 - 1.4) | 1.5 (.6 - 2.2) |
| | Cz | .9 (.1 - 1.6) | .3 (- .4 - 1.0) | 1.2 (.4 - 1.9) |
| Electrode X Group (crct) | Fz | 1.2 (.4 - 1.9) | .1 (.6 - .8) | 1.4 (.6 - 2.2) |
| | Fcz | 1.1 (.3 - 1.8) | .1 (.6 - .8) | 1.2 (.4 - 1.9) |
| | Cz | 1.0 (.2 - 1.7) | .1 (.6 - .8) | 1.2 (.4 - 2.0) |
| | Cpz | .9 (.1 - 1.6) | .1 (.-6 - .8) | .9 (.1 - 1.6) |
| Electrode X Group (wrg) | Fz | .8 (.1 - 1.5) | .8 (0.0 - 1.5) | 1.5 (.7 - 2.2) |
| | 750-900ms | | | |
| Electrode X Group (all) | Fz | .9 (.2 - 1.7) | .1 (-6 - .8) | 1.0 (.3 - 1.8) |
| Electrode X Group (crct) | Fz | .9 (.2 - 1.6) | .0 (-.7 - .7) | .9 ( .2 - 1.6) |
| Electrode X Group (wrg) | Fz | .6 (-.1 - 1.3) | .4 (⁻.3 - 1.1) | .9 (- 1.6 - ⁻.2) |

| Overall ES from Old/New task: | | 1.0 | .3 | 1.5 |
|---|---|---|---|---|

**Frontal vs. Posterior Electrodes**

| | 400-650ms | | | |
|---|---|---|---|---|
| Frontal/Posterior X Group (all)  F | | .9 (.2 - 1.7) | .3 (-.5 - 1.0) | 1.2 (.4 – 2.0) |
| Frontal/Posterior X Group (crct) F | | 1.0 (.2 - 1.7) | .1 (-.7 - .8) | 1.0 (.3 - 1.8) |
| | 650-750ms | | | |
| Frontal/Posterior X Group (all)  F | | 1.1 (.3 - 1.8) | .7 (.1 - 1.4) | 1.6 (.7 - 2.4) |
| Frontal/Posterior X Group (crct) F | | 1.1 (.3 - 1.8) | .4 (-.4 - 1.10) | 1.5 (.6 - 2.3) |
| Frontal/Posterior X Group (wrg) F | | .8 (.1 - 1.5) | .8 (.0 - 1.5) | 1.5 (.6 - 2.3) |
| | 750-900ms | | | |
| Frontal/Posterior X Group (all)  F | | .8 (.0 - 1.5) | .5 (-.3 - 1.2) | 1.1 (.3 - 1.9) |
| Frontal/Posterior X Group (crct) F | | .8 (.1 - 1.5) | .2 (-.5 - .92) | .9 (.1 - 1.7) |
| Frontal/Posterior X Group (wrg) F | | .6 (.2 - 1.3) | .7 (.1 - 1.40) | 1.1 (.3 - 1.8) |

| Overall ES frontal electrodes: | | .9 | .4 | 1.2 |
|---|---|---|---|---|

*Note*. HON = Honest Responders; MAL = Malingering Responders; TBI = Traumatic Brain Injury group; crct = correct responses; wrg = wrong responses; F = frontal electrodes. ES value represents absolute value.

Table 12
*Effect Size Data in Cohen's d (Significant Results for the Repetition Priming Task)*

| Electrodes Analyzed | | | Groups Compared: ES (95% CI) | | |
|---|---|---|---|---|---|
| Variable | Interval | | HON v MAL | MAL v TBI | HON v TBI |

**Midline Electrodes**

| | | | | | |
|---|---|---|---|---|---|
| | 0-200ms | | | | |
| Rep X Group (wrg) | 2nd | | .7 (.0 - 1.4) | 1.0 (.2 - 1.7) | .3 (-.5 - 1.0) |
| | 400-600ms | | | | |
| Electrode X Group (all) | | Fz | .8 (.0 - 1.5) | .3 (-.4 - 1.0) | 1.2 (.4 – 2.0) |
| | | Fcz | .9 (.2 - 1.6) | .1 (-.6 - .8) | 1.3 (.5 - 2.0) |
| | | Cz | .9 (.1 - 1.6) | .2 (-.5 - .9) | 1.2 (.4 - 2.0) |
| Electrode X Group (crct) | | Fz | .7 (.0 - 1.4) | .2 (-5 - .9) | 1.1 (.3 - 1.8) |
| | | Fcz | .9 (.1 - 1.6) | .0 (-.7 - .7) | 1.1 (.3 - 1.8) |
| | | Cz | .7 (.0 - 1.5) | .1 (-.6 - .8) | 1.0 (.3 - 1.8) |
| | 600-900ms | | | | |
| Electrode X Group (all) | | Fz | .8 (.0 - 1.5) | .4 (-.4 - 1.1) | 1.4 (.5 - 2.1) |
| | | Fcz | .9 (.2 - 1.6) | .1 (-.6 - .8) | 1.1 (.3 - 1.8) |
| Electrode X Group (crct) | | Fz | .7 (.0 - 1.5) | .2 (-.5 - .9) | 1.2 (.4 - 1.9) |
| | Overall ES: | | .8 | .3 | 1.1 |

**Frontal vs. Posterior Electrodes**

| | | | | | |
|---|---|---|---|---|---|
| | 400-600ms | | | | |
| Frontal/Posterior X Group (all) | F | | .9 (.1 - 1.6) | .7 (-.1 - 1.4) | 1.6 (.8 - 2.4) |
| Nov X Rep X Group (all) | Old/2nd | | .5 (-.2 - 1.2) | .4 (-.3 - 1.2) | 1.0 (.2 - 1.7) |
| | 3rd | | .8 (.1 - 1.5) | .5 (-.3 – 1.2) | 1.3 (.2 - 2.0) |
| | New/2nd | | .6 (-.2 - 1.2) | .5 (-.3 - 1.2) | 1.0 (.2 - 1.7) |
| | 3rd | | 1.0 (.2 - 1.7) | .1 (-.6 - .8) | .9 (.1 - 1.6) |
| Frontal/Posterior X Group (crct) | F | | .8 (.1 - 1.5) | .6 (-.2 - 1.3) | 1.5 (.7 - 2.2) |
| Frontal/Posterior X Group (wrg) | F | | .8 (.0 - 1.6) | .6 (-.2 - 1.3) | 1.5 (.7 - 2.3) |
| | 600-900ms | | | | |
| Frontal/Posterior X Group (all) | F | | .8 (.0 - 1.5) | .7 (-.1 - 1.4) | 1.4 (.6 - 2.2) |
| Frontal/Posterior X Group (crct) | F | | .7 (.1 - 1.5) | .5 (-.2 - 1.2) | 1.3 (.5 - 2.0) |
| Frontal/Posterior X Group (wrg) | F | | .8 (.0 - 1.6) | .7 (-.1 - 1.4) | 1.7 (.8 - 2.5) |
| | Overall ES: | | .9 | .5 | 1.5 |

*Note*. HON = Honest Responders; MAL = Malingering Responders; TBI = Traumatic Brain Injury group; crct = correct responses; wrg = wrong responses; Nov = Novelty; Rep = Repetition; 2nd = 2nd presentation of item; 3rd = 3rd presentation of an item; F = frontal electrodes. ES value represents absolute value.

*3.24 Classification Statistics for ERP Data*

The differences between the three groups for ERP responses were compared in order to choose the best variables to classify MAL and honest (clinical and non-clinical) groups. Table 13 presents selected ERP variables and the corresponding cutting scores that yielded the highest overall hit rates for each variable. For the Repetition Priming task, the variable with the largest ES was the mean ERP amplitude for the 2nd presentation of an item during the early interval of 0-200ms using wrong responses. The classification data for this variable were acceptable the hit-rate for detecting malingering for HON was 74% and increased for TBI to 78%. While this variable had the highest hit rate when comparing HON and MAL, this comparison is less clinically relevant than TBI vs. MAL (the cost of misidentifying an individual with true brain impairment is high). In addition, the group differences seen in this variable are isolated with no supporting data in later intervals. For both tasks ERP difference scores appeared to be the best way to classify malingering responders, which may be due to the fact that the malingering group was less likely to show differential responses (old/new items; frontal/posterior variation) compared to HON and TBI. As mentioned in an earlier section, the data suggest that HON and TBI had distinct responses to old vs. new items in the Old/New task. The ability to visually inspect ERP waves to detect MNCD was only moderate. Since the classification statistics for that fast "eye-balling" method of detecting malingering were low (65% and 70% for HON and TBI, respectively), a smaller time interval (400-650 ms) was used in a later test. The difference score was calculated using statistical analysis in order to increase the classification hit-rate. As a result, the differences in ERP mean amplitude for new and old items (all responses) for the interval between 400–650 ms was chosen as one ERP variable that might be useful in detecting MNCD. Using this revised difference score, the hit-rate for detecting malingering decreased for HON (50%) and increased for TBI (77%). While this Old/New difference score improved hit rates between TBI and MAL groups, the variable was not considered the most effective way to detect MNCD due to the unacceptably low hit rate when comparing HON and MAL.

Results from the Old/New task also indicated that the differences between frontal and posterior electrodes were significant for HON and TBI but not different in MAL. The ES results suggested that the mean ERP amplitude for the frontal electrodes during the Old/New task may differentiate between TBI and MAL (Cohen's $d$ = .8, as mentioned earlier, Cohen's $d$ = .8 is considered a large ES). The best variable to use as a frontal–posterior difference score to identify malingering was picked based on the ES for mean differences among groups. When examining ES for frontal vs. posterior responses, the ERP mean amplitude for frontal electrodes incorrect responses during the later interval of 650-750 ms (F-P difference score) was determined to be the best ERP variable to detect MNCD. The mean differences between HON/TBI vs. MAL were not significant for any posterior comparisons; however there were group differences for frontal electrodes. Since the ES for the frontal electrodes showed the largest difference between the MAL and TBI group, this variable was determined to be a good candidate to detect MNCD. Although incorrect trials were general low for the three groups, the number of trials for each participant for this analysis yielded approximately 20-35 trials because both old and new variables were included. In addition, the differences in the mean ERP amplitude for the three groups was consistent across types of responses (all, correct, and incorrect), most intervals, and both memory tasks which indicates this difference is not random.

The Frontal-Posterior (F-P) difference score from the Old/New task produced a 58% hit rate in classifying HON and MAL responders. The F-P difference had a fairly

Table 13

*Classification Statistics for ERP Variables (Base Rate for Simulation Malingering=50%)*

| Variable | **Groups Compared** | Sensitivity | Specificity | Hit Rate | NPP | PPP |
|---|---|---|---|---|---|---|
| | **HON (non-clinical) vs. MAL** | | | | | |
| <u>Old/New task</u> | | | | | | |
| N-O wave dif (< -.75 = mal) | | 75.0 | 37.5 | 50.3 | 74.4 | 38.2 |
| F-P wave dif-wrong (< 2.0 = mal) | | 46.7 | 69.2 | 58.0 | 56.5 | 60.3 |
| electrode Fz-visual inspection of old & new waves (not malingering = old > new; 300-700ms) | | 66.7 | 62.5 | 64.6 | 65.2 | 64.0 |
| <u>Repetition Priming</u> | | | | | | |
| Midline electrodes mean ERP amplitude for 2nd presentation-wrong (< .25 = mal) | | | | | | |
| | | 60.0 | 81.3 | 74.1 | 67.0 | 76.2 |
| | **TBI (clinical honest) vs. MAL** | | | | | |
| <u>Old/New task</u> | | | | | | |
| N-O wave dif (< -.75 = mal) | | 87.5 | 66.7 | 77.1 | 84.2 | 72.4 |
| F-P wave dif-wrong (< 2.0 = mal) | | 80.0 | 92.9 | 86.5 | 82.3 | 91.9 |
| electrode Fz-visual inspection of old & new waves (not malingering = old > new; 300-700ms) | | 66.7 | 73.3 | 69.7 | 75.3 | 64.3 |
| <u>Repetition Priming</u> | | | | | | |
| Midline electrodes mean ERP amplitude for 2nd presentation-wrong (< .25 = mal) | | | | | | |
| | | 60.0 | 86.7 | 77.6 | 70.0 | 80.9 |

*Note*.  HON = Honest Responders; MAL = Malingering Responders; TBI = Traumatic Brain Injury group.  N-O wave dif = new – old wave difference; F-P wave dif = new – old wave difference; NPP = Negative Predictive Power; PPP = Positive Predictive Power.

high hit rate of 87% when differentiating between the TBI and MAL groups. Overall, for most of the ERP variables, the hit rates when comparing TBI vs. MAL were higher than when comparing HON vs. MAL. Since the TBI vs. MAL comparison is considered a more important comparison to correctly identify MNCD without misdiagnosing true brain injury, the F-P difference was determined to be the best ERP biomarker to predict malingering status.

*4. Logistic Regression*

Logistic regression is a statistical analysis that can help identify successful variables that aid in the prediction of a dichotomous outcome (e.g., honest vs. malingering). This type of analysis offers one way to compare the unique predictive power of different methods of detecting MNCD. There are several important statistics used to interpret logistic regression. First, the proposed model is tested against the null (no variables in the model, only a constant). If the model indicates that the variables significantly predict the probability of malingering status, the chi-square statistic will be significant ($p < .05$). For this study, the outcome variable is malingering (coded as 1) vs. honest (coded as 0). Successful accuracy, RT, and ERP variables mentioned in earlier sections will be entered into the model as predictive variables that can be continuous or dichotomous. A regression coefficient ($\beta$) determines the relationship between the outcome and predictor variables. When $\beta = 0$ there is no relationship between the two variables. If the regression coefficient is positive, it means that there is a positive relationship between the outcome and predictor (e.g., as RT increases the probability of being classified as malingering increases). Conversely, if the regression coefficient is negative the relationship is negative, (e.g., as accuracy increases the probability of being classified as malingering decreases). The odds ratio (OR) is the odds of being classified as malingering for a given predictor variable. For a dichotomous variable (e.g., for gender, male = 1 and female = 0), the OR indicates the likelihood of being classified as malingering if the participant is male (the variable coded as "1") compared to female. For continuous variables, the odds ratio signifies the change in odds for participants to be classified as malingering for each unit increase in the predictor variable (e.g., if accuracy had an OR = 2.0, for every additional unit of accuracy the likelihood of being classified as malingering increased 2 times). Finally, there is a corresponding hit rate for each variable that significantly entered into the model.

The accuracy of both the TOMM-C Trial 3 and old items in the Old/New task had classification hit rates that were perfect (i.e., TOMM) or near perfect (i.e., Old/New task). Thus, that meant that for technical reasons, these two variables could not be used in a logistic regression because the maximum likelihood estimates (such as the Wald test) can not be calculated when there is no overlap in the distribution of the variables between the two outcome groups (honest vs. malingering). While the success of these two accuracy variables in detecting malingering is duly noted, it is also important to examine additional malingering detection variables since the field of detecting MNCD must stay ahead of individuals who seek information regarding current malingering tests. Currently, accuracy is the variable used in malingering tests to detect MNCD, and newer variables (i.e., RT and ERP) may be more effective in detecting malingerers.

Since the best TOMM and Old/New accuracy variables could not be used in the logistic regression, the best accuracy score for the Repetition Priming task was applied instead. The accuracy for the 1st repetition of an item was chosen since it had a low rate of false positives when differentiating between TBI and MAL, and the hit rates for the HON vs. MAL comparison was moderately successful as well (approximately 78%). As

mentioned in the RT results section, the Trial 3 RT for correct items in the TOMM-C was the best RT variable with a hit rate of 78% and 71% for HON and TBI respectively. The best ERP predictor was determined to be the difference score between frontal and posterior ERP mean amplitude from 650-750 ms to items answered incorrectly in the Old/New task. The following regression analyses compared the best accuracy, RT, and ERP variables in order to determine if the newer variables of RT and ERP increase classification of malingering.

A forward conditional stepwise logistic regression analysis was performed on malingering status and the best accuracy, RT and ERP predictors. The variables entered into the model were Repetition Priming accuracy for the 1st repetition of items (RP 1st accuracy), the Trial 3 RT for correct items in the TOMM-C, and the difference score between frontal and posterior ERP mean amplitude from 650-750ms to items answered incorrectly in the Old/New task (F-P difference score). Based on the sample size for the groups examined in the analyses ($N$'s = 27-28), only three variables could legitimately be used as predictor variable (one for each method). The number of predictors should not exceed the ratio of approximately 10 to 1 or the stability of the statistical analysis is questionable. It should be noted that the RP accuracy variable was converted into a dichotomous variable with an 84% cutting score, because the skewed nature of the data as a continuous variable was not compatible with the regression analysis. Thus, the data are accurately represented in a logistic regression analysis by converting them into a dichotomous variable. The 84% cutting score was used to divide the variable into two groups because it was the most successful cutting score (> 84% coded as 1, < 84% coded as 0). Logistic regression analysis was used to see what type of method of detecting malingering was the most successful in predicting malingering status, specifically if RT and/or ERP added any additional predictive power.

There were two conceptually important comparisons for this experiment; HON (non-clinical, $N = 16$) vs. MAL ($N = 16$) and TBI (clinical honest, $N = 15$) vs. MAL that were undertaken to determine if certain variables performed differently when detecting malingering status compared to different types of honest responding groups (HON or TBI). For each comparison the RP 1st accuracy score, the Trial 3 RT for correct items in the TOMM-C, and the F-P difference score (ERP) were entered in a stepwise regression. Table 14 shows the logistic regression coefficient ($\beta$), Wald test, OR, and hit rates for variables that significantly entered into the logistic regression for both comparisons. For the first comparison (HON vs. MAL), a test of the full model versus a model with only the constant was statistically significant, $\chi^2(2, N = 27) = 17.71, p < .0001$. The RP 1st accuracy was the only variable that entered into the model and predicted 85% of HON and 79% of malingerers for an overall 82% hit rate. The inverted OR indicated that participants who had accuracy scores > 84% the probability of being classified as malingering was 25 times less than participants with accuracy scores < 84% (OR = .04; 1.0/.04 = 25; $p < .01$). In classifying non-clinical honest and MAL responders, the RP 1st accuracy variable was the only significant predictor of malingering status.

For the second analysis, comparing the TBI (clinical honest) and MAL, a test of the full model versus a model with only the constant was statistically significant, $\chi^2(2, N = 28) = 20.16, p < .0001$. The F-P difference score entered into the model and predicted 93% of TBI and 79% of malingerers for an overall 86% hit rate. The odds ratio for the F–P difference score indicated that for every unit (1 mV) increase in the F–P mean amplitude the odds of being classified as a malingerer increased 2.12 times ($p < .01$). The results suggest that when classifying MAL and TBI honest the F-P difference score

68

Table 14

*Logistic Regression of Malingering Status as a Function of Accuracy, RT, & ERP*

*Variables*

| Significant Variables | | Wald | **Odds** | 95% CI for Odds Ratio | | | |
|---|---|---|---|---|---|---|---|
| **Groups Compared** | *β* | Test | **Ratio** | *Lower* | *Upper* | *p* | **HR** |
| **HON (non-clinical) vs. MAL** | | | | | | | |
| RP task-1st item accuracy | -3.17 | 7.37 | **.04** | .00 | .41 | .01 | **82%** |
| Constant | -1.36 | .48 | .26 | | | .49 | |
| **TBI vs. MAL** | | | | | | | |
| F-P dif wave | .75 | 6.84 | **2.12** | 1.21 | 3.73 | .01 | **86%** |
| Constant | -2.76 | 2.76 | .10 | | | .06 | |

*Note.* HON = Honest Responders; MAL = Malingering Responders; TBI = Traumatic Brain Injury group; RP = Repetition Priming task; ac. = accuracy (% correct); T3 RT = Trial 3 reaction time; *df* = 1 for all variables; *β* = regression coefficient (+ value denotes positive relationship with malingering, - value denotes negative relationship); Odds Ratio = the odds of the variable predicting malingering status (ratios for RP variable [dichotomous] reflect odds of classification of malingering if accuracy > 84%); F-P dif wave = difference in mean amplitude for frontal vs. posterior electrodes incorrect responses; HR = hit rate for variable in current step and all prior steps.

predicted the status of participants such that as the ERP mean amplitude for the frontal electrodes increased compared to the posterior, the odds of being classified as malingering more than doubled. Overall, the logistic regression results indicated that the two honest responding groups (HON and TBI) were differentiated from MAL using different methods. Accuracy was the only significant predictor when comparing HON vs. MAL, while the ERP variable was the only significant predictor when comparing TBI vs. MAL. It should be noted that the F-P difference score led to 93% specificity and an 86% hit rate which approaches the TOMM accuracy classification rates.

## 5. ERP Results Showing HON vs. TBI Differences

The difference between HON and TBI across tasks indicated that normals (non-brain injured) and a brain injured population (TBI) look very different on a measure of brain activity. The ERP analyses showed consistent group differences such that healthy, honest responders had significantly larger mean ERP amplitudes compared to TBI. This comparison was not directly relevant to the focus of this study (i.e., detecting MNCD); however it is significant for future studies attempting to examine the ability of ERPs to detect MNCD. While earlier sections note differences in ERPs for malingering and TBI groups, there are also larger differences between honest non-brain injured people and TBI seen in ERP data. If these data are not reported, inaccurate criteria could be developed that would classify honest responders as malingering since the overall means of non-brain injured groups (both malingering and honest) have larger ERPs than TBI. Another benefit of noting the HON vs. TBI comparison reveals information about the effect of brain injury on cognitive processes that have been well documented in healthy individuals, but less so in a population with TBI. Understanding the effect of brain injury on ERPs is essential if this method is to become a validated method to detect MNCD. If the effect of brain injury is not documented and compared to both honest, non-brain injured and feigners the risk of misdiagnosis in brain injured individuals as malingering is high.

Both memory tasks examined in this experiment showed reliable group differences such that healthy HON had larger ERP responses than TBI. For the Old/New task, the significant group differences occurred between HON and both the TBI and MAL groups. The largest group differences typically occurred during the later time intervals 400-900 ms in the frontal midline area (generally, Fz-Cz). Overall, ES data of group means for frontal electrodes Fz–Cz during 400–900 ms (P300 or LPC) support the idea that ERPs in the Old/New task reveal a large difference (mean Cohen's $d = 1.5$). The difference was also large when comparing HON vs. TBI in mean ERP amplitude for the average of the frontal electrodes the HON and TBI groups (mean Cohen's $d = 1.2$). This difference was also seen when examining group differences from the frontal electrodes in the Repetition Priming task (mean Cohen's $d = 1.5$). In addition, HON had significantly larger ERPs compared to TBI (mean Cohen's $d = 1.1$) for the frontal midline electrodes, particularly in the late components from 400-900 ms. The present results have demonstrated that brain injury greatly decreased mean ERP amplitudes across memory tasks. This result makes the specific differences in MAL vs. TBI (such as using within group difference scores) noted in earlier sections especially important since only using mean group differences in ERP would not detect malingerers since HON had the largest ERP amplitude, TBI had the smallest, and MAL was in between both groups.

Chapter Four
Discussion and Conclusions

The aim of this study was to determine the effectiveness of using two new methods to detect malingered neurocognitive deficit (MNCD), i.e., RT and ERP measures. The longer the standard malingering tests that use accuracy to detect MNCD (e.g., TOMM) are used in clinical practice and studied in academic settings, the greater the likelihood of people evading detection due to their knowledge of existing ways to detect malingering. For this reason two additional methods, RT and ERP, were examined for future use in the detection of MNCD. These new measures were examined against the established method of accuracy on standard malingering tests such as the computerized version of the TOMM (TOMM-C).

*Accuracy*

The TOMM-C accuracy results were consistent with prior research supporting its use as a malingering test (Tombaugh, 1997; Rees et al., 1998). The current computer version of the TOMM results produced high classification rates where HON and TBI near perfect (99-100%) accuracy while MAL had much lower accuracy (61-64%). More importantly, hit rates for detecting malingering were 100% using the TOMM Trials 2 and 3 accuracy variables.

The present study also measured accuracy for a computerized Old/New recognition task. The accuracy for studied items from the Old/New task was successful in differentiating between honest and TBI vs. malingering responders with hit rates of 88% and 74%, respectively. In addition, the accuracy variables from the TOMM and Old/New task had reasonable success in distinguishing honest and malingering responders and could not be used in a logistic regression analysis. Therefore, the role of a new accuracy variable from the Repetition Priming task was also examined in a logistic regression comparing the best accuracy, RT and ERP variables.

When logistic regression was employed to evaluate the HON vs. MAL comparison, the accuracy variable from the Repetition Priming task was the only significant predictor of malingering status. The RP accuracy for the 1$^{st}$ presentation of an item resulted in an 82% hit rate. As mentioned above, the standard malingering tests use accuracy to detect MNCD in clinical practice and have been studied in academic settings for a long time. The likelihood of people evading detection increases as more information on existing tests is available on sites such as the internet or academic journals. For this reason developing accuracy measures in new tasks and examining additional methods such as RT and ERP are pivotal for continued success in identifying individuals who feign neurocognitive deficit.

*RT*

RT data was calculated for the TOMM-C, Old/New task, and Repetition Priming task. When the task allowed the participant to take as long as necessary to respond (as in the TOMM-C test format), RTs from Trial 3 to correct responses obtained the highest hit rates for a RT variable (hit rates of 71%-78%). Employing a cutting score of > 1575 ms had a sensitivity of 63% and a specificity of 80% for MAL vs. TBI groups and a sensitivity of 63% and a specificity of 94% for MAL vs. HON. This result may have occurred because TBI had enough time to respond as accurately as possible, and MAL has more time to "fake" their response. In contrast, when the task had a defined response interval such as in the Old/New and Repetition Priming tasks where responders had approximately 1.5 seconds (or 1500ms) to respond, the primary differences occurred between the HON and TBI groups such that TBI had significantly longer RTs than HON.

The relative success of RTs obtained from a task with unlimited time to respond is contrary to prior research from Rose and colleagues (1995) who found that a task with no time limit produced RTs for a TBI group that were slower than for malingerers and much longer in duration (approximately 4 seconds). In future studies, tasks attempting to use a RT variable to detect MAL should have unlimited time (or a long enough time to respond such as 3-4 seconds) similar to the TOMM-C test used in this experiment.

A pattern of RTs has also shown promise as a detection strategy as seen in the results when comparing RTs to old vs. new items in the Old/New task. Both HON and TBI produced the classic "old/new effect" with slower RTs to new pictures compared to old pictures, while MAL had no difference. These results replicate prior research findings that there may be some validity in examining tasks with different RT patterns for different items to detect feigning cognitive impairment (van Gorp et al., 1999; Osimani et al., 1997). Although the old-new RT difference score was not as successful as the TOMM T3 RT (hit rate 71% compared to 78%), there is evidence that this variable has some ability to detect MNCD. In a future study, examining an old-new RT difference when participants have a longer time to response may further differentiate between honest (HON and TBI) and malingering responders.

*ERP*

The ERP results across tasks in this study indicated that the ERP mean amplitude was largest for honestly responding healthy adults, and furthermore the effect of malingering in non-brain injured participants appears to reduce ERP mean amplitude so these responders look similar to participants with TBI (the mean ERP amplitude for MAL was only slightly larger than TBI). These results are similar to prior research by Ellwanger et al. (1996), and contrary to findings of Ellwanger et al. (1999) which indicated that malingerers had a larger P300 than honest responders. In addition, the findings from this study suggest that MAL and TBI look similar when examining ERP mean amplitude across all analyses (midline, hemisphere, frontal vs. posterior) where prior studies predominantly examined only the P300 component at electrode Pz.

The present results with 32-channel recording have shown that several ERP measures can be used to differentiate among the groups. First, during the Old/New task, ERPs from the front half of the scalp revealed a significant difference among the groups for the Old/New task from 400-900 ms. HON had a larger ERP wave segment in frontal electrodes compared to posterior, while TBI had larger mean amplitude in the posterior compared to frontal electrodes. There was no significant difference in ERPs of frontal versus posterior electrodes for MAL. This is similar to prior reports that across three midline electrodes (Fz, Cz, and Pz) malingerers did not show differences found in an honest responding group (Rosenfeld et al., 1999). The distribution of activity in the non-brain injured, honestly responding person appears to occur primarily in the frontal region, and in TBI the primary activity is in the posterior part of the brain (perhaps compensation for damage as a result of a head injury). The lack of difference for MAL may be due to a focus on manipulating responses that may reduce the activity in the frontal electrodes. This can be seen in comparisons indicating HON had significantly larger mean amplitude in the frontal electrodes than MAL. The reduced mean ERP amplitude in MAL may be due to several factors. Kok (2001) gives several factors that affect P300 amplitude which may illustrate what is occurring in the malingering group to lead to reduced mean amplitude in the frontal electrodes. Reduced ERP amplitude may have been related to decreased attention to the primary task due to a secondary task of malingering. Task difficulty also leads to a decreased P300, and for MAL the additional aspect of

malingering may have increased the complexity of the task.  Finally reduced confidence in one's performance is related to a decrease in P300.  Malingerers may not be sure if they are answering in a way that will appear as if they have brain damage, and their confidence suffers during the task compared to HON.

Second, when examining the classic peaks for the Old/New task, the malingerers did not show the classic "old/new effect" in contrast to the honest responders (HON and to a lesser extent TBI).  MAL did not have any peak differences between old and new items.  HON had larger P200 peak amplitude for old items compared to new items for the frontal and central electrodes Fz, Fcz, Cz, Cpz, and Pz.  TBI had larger peak amplitude for old compared to new items as the posterior electrode, Oz.  Honest responders (HON and TBI) differentiate between old and new items, but this differentiation is absent in malingerers.  Malingering appears to reduce inter-item ERP differences.  Similar to the finding from the frontal vs. posterior analyses, malingering reduced ERP differences seen in the two honest responding groups.  The P300 was the predominant component examined in prior research using ERP to detect MNCD.  For the Old/New task, HON had the classic "old/new effect" response (mean amplitude for old items > than new), and although the difference for TBI did not reach significance the trend was in the right direction.  Once again malingerers showed no difference between old and new items.  In a review of existing studies in this area, Rosenfeld and Ellwanger (1999) theorized that the lack of difference in ERPs for different items in malingerers may be due to the increased workload when manipulating responses.  In the present study, malingerers did not show variation in ERPs to items in the Old/New task that were seen in both intact and TBI responders performing honestly.  When comparing TBI (clinical honest) and MAL the ERP variable of the F-P difference score (86% hit rate) was a significant predictor. The ERP variable had unique classification accuracy when used to differentiate TBI (clinical honest) and MAL.  For the two new methods examined in this study, only ERP added unique prediction power when comparing TBI vs. MAL.  ERP data have been used in prior research to attempt to find a physiological marker for malingering (Rosenfeld et al., 1995; Rosenfeld et al., 1998; Ellwanger et al., 1999; Tardif et al., 2002; etc.).

*Importance of a Clinical Control Group*

Most of the studies examined the differences in responding honestly compared to malingering without a clinical control group.  The current study illustrates the importance of having a clinical control when beginning to explore new methods to detect MNCD. While there were large and consistent differences in ERPs between the HON and MAL groups in this study, TBI generally appeared similar to MAL.  If characteristics that distinguish malingering from non-clinical honest were used to classify feigning, many truly impaired people would be misdiagnosed as malingering.  Yet, an ERP variable was successful in adding classification accuracy when comparing a truly brain injured group with a MAL group.  ERP may be more sensitive in identifying brain injury rather than malingering and closer inspection of tasks may lead to physiological markers that can identify the presence of TBI (and thus rule out malingering if the marker is present). Based on the results from this study a task that requires frontal manipulation such as the Old/New task could be successful in differentiating between honest and malingering responders because TBI and HON show frontal vs. posterior differences while MAL does not.

*Limitations of the Current Study*

Limitations of the current study are the lack of generalizability of the simulation setting where no realistic system of reward or punishment was implemented for the

malingering condition that would mirror a real forensic setting.  In the future, studies attempting use RT and/or ERP methods to detect MNCD could offer participants instructed to malinger an extra incentive for successful malingering and other recommendations made by Rogers (1997).  For the Old/New and Repetition Priming tasks the interval to respond was very short (approximately 1.5 s), and this may have reduced the difference between the MAL and TBI groups because when forced to respond very quickly MAL participants were able to adjust and responded faster than participants with cognitive impairment; however it is important to note that for ERP research, extending the interval for participants to respond often decreases the number of usable events to average, especially for special populations such as participants with brain injury.  Another limitation of this study is the lack of ability to correlate the area of brain injury with specific ERP results (such as frontal vs. posterior differences).  The participants that made up the TBI group had documented brain injury, but the injuries were widespread across the brain.  This may have reduced the statements that can be made about how malingerers perform differently than TBI on the implicit (Repetition Priming) and explicit (Old/New) memory tasks used in this study.  A future study examining participants with damage in a specific area of the brain may be able to make more direct statements about how TBI and malingering are different.

*Future Directions*

   While the hit rate for ERP in detecting MNCD was not as high as the traditional paper and pencil malingering test, there were clear group differences between the HON, TBI and MAL groups which warrant further investigation.  Future studies could determine the source of the reduced ERP mean amplitude in malingering responders compared to honest responders, specifically in the frontal electrodes.  There are several ways the nature of the reduction could be examined such as having a feigning group instructed to pay attention to the study phase of items and a group instructed not to pay attention to items enabling the effect of reduced attention on ERP results.  Malingering participants could also be questioned about their malingering strategy, especially what they were thinking about as they responded to test items in order to find common strategies that may be reducing the mean amplitude when someone feigns.  A future study could compare different strategies when people malinger and the effect on ERP amplitude in order to see if one aspect of malingering is responsible for the decrease in ERP amplitude.  In the current study both the explicit memory (Old/New) task and the implicit memory task of Repetition Priming appeared to show similar patterns of group differences.  Both tasks elicited more frontal ERP mean amplitude in honest responders (F > P) and helps differentiate between a brain injured group that has larger posterior mean amplitude (P > F) and a MAL group with no significant differences across the scalp (F = P).  Perhaps a future study could include a different task that requires frontal resources such as a matching to sample task in order to see if this pattern of activity for the three groups remains when tested with another task requiring frontal manipulation.  In addition, factors that affect mean ERP amplitude mentioned earlier such as decreased attention, task difficult, and confidence in performance could also have played a role in the reduced mean amplitude for the TBI as well as MAL.  A matching to sample task with a longer response time may increase the mean amplitude for TBI (since the task should be easier); however, MAL will still be malingering and have reduced mean amplitude.

*Conclusions*

   In conclusion, this study is the first study that compared three groups in a

malingering paradigm using multiple EEG electrodes.  Accuracy, RT, and ERP variables from a standard malingering test and two new cognitive tasks were examined to see how each measure detected MNCD.  The standard malingering test (TOMM) replicated prior data supporting its use to detect MNCD.  Accuracy measures from two new tasks (Old/New and Repetition Priming) were also successful in classifying malingering.  The newer RT measure obtained from a task with no set response time also appears to have some success in detecting MNCD.

Although the hit rates using ERP to detect MNCD were not as high as those for the standard malingering test, the current results suggest that there were group differences between HON, TBI, and MAL responders detected in ERP data.  Specifically one ERP variable was able to offer unique prediction power when differentiating between TBI and MAL groups.  Furthermore, the ERP results support the recommendations of Rogers (1997) that a clinical honest group (TBI) is necessary in evaluating new malingering detection methods.  The inclusion of a clinical comparison group is especially important when examining the ability of ERP to detect MNCD since malingerers often had ERP mean amplitudes that were no different than TBI or in between the honest and TBI groups.  Basing a detection strategy on ERP data between only intact honest and malingering responders could lead to misclassification of truly brain injured individuals as malingering.  Direct comparison among the three groups also helps to identify the subtle differences due to malingering, which allows for better classification rates.  The current study demonstrates the potential clinical use for ERP methods in detecting MNCD.

Appendix A

Instructions for Participants

Instructions for HON and TBI participants:

You will be asked to answer questions on different tests that were designed to test your cognitive function.  Please give your best effort on each of the tests you take.  Do you have any questions about these instructions?


Instructions for MAL participants:

For the next 3 tests we would like for you to pretend that you are going through an evaluation after having a car accident.  You want to appear as if you have suffered brain damage after the accident in order to increase your monetary reward from your insurance company.  If your test results indicate that you have brain damage than you will receive a great deal of money; however, you do not want to perform so badly that the tester knows you are faking brain injury.  Do you have any questions about these instructions?

Appendix B

Instructions for the Three Tasks

INSTRUCTIONS FOR 1<sup>ST</sup> TASK

For this task you will see 50 pictures on the computer screen. Try and remember these pictures as best you can because afterwards 2 pictures will appear on the computer screen and I want you to choose which picture you remember from the pictures you memorized. Press the "→" key or the "←" to highlight the picture you remember and press "Spacebar" to go to the next item. Even if you are not sure of the answer, give your best guess.

INSTRUCTIONS FOR MEMORY (2<sup>ND</sup>) TASK

1<sup>st</sup> Part - For the next task you will see a new series of objects. Try to memorize each object as best as you can. The best way to remember something is to think about how it relates to yourself or someone you know. You will be tested on these pictures later on. 2<sup>nd</sup> Part - For this next task you will see a picture on the screen. Press the right, blue "MEM" key if it is a picture you remember from the latest series of pictures you memorized, and press the left, yellow "NOT MEM" button if it is a picture you have never seen before.

INSTRUCTIONS FOR THE CATEGORY (3<sup>RD</sup>) TASK

Several objects will be presented. Your job is to determine whether each object is man-made or not. Thus, a newspaper is man-made while a stick or tree branch is not. If the object is man-made, then press the right, blue "MAN-MADE" key. If the object is NOT man-made, then press the left, yellow "NOT MAN-MADE" key. Some items may be considered both man-made and not man-made, just go with your first impression.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., Text Revision). Washington, DC: Author.

Bashore, T. R., & Rapp, P. E. (1993). Are there alternatives to traditional polygraph procedures? *Psychological Bulletin, 113*(1), 3-22.

Berry, D. T. R., Baer, R. A., Rinaldo, J. C., & Wetter, M. W. (2002). Assessment of malingering. In J. Butcher (Ed.), *Clinical personality assessment: Practical approaches*. New York: Oxford University Press.

Binder, L. M., & Willis, S. C. (1991). Assessment of motivation after financially compensable minor head trauma. *Psychological Assessment: A Journal of Clinical and Consulting Psychology, 3*, 175-181.

Brown, M.W., & Aggleton, J. P. (2001). Recognition memory: What are the roles of perirhinal cortex and hippocampus? *National Review Neuroscience, 2*, 51-61.

Cycowicz, Y. M., Friedman, D., Rothstein, M., Snodgrass, J. G. (1997). Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology, 65*(2), 171-237.

Cohen, J. (1998). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Curran, T., & Cleary, A. M. (2003). Using ERPs to dissociate recollection from familiarity in picture recognition. *Cognitive Brain Research, 15*, 191-205.

Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences USA, 93*, 13494-13499.

Ellwanger, J., Rosenfeld, J. P., & Sweet, J. J. (1997). P300 event-related brain potential as an index of recognition response to autobiographical and recently learned information in closed-head injury patients. *The Clinical Neuropsychologist, 11*, 428-432.

Ellwanger, J., Rosenfeld, J. P., Sweet, J. J., & Bhatt, M. (1996). Detecting simulated amnesia for autobiographical and recently learned information using the P300 event-related potential. *International Journal of Psychophysiology, 23*, 9-23.

Ellwanger, J., Tenhula, W. N., Rosenfeld, J. P., & Sweet, J. J. (1999). Identifying simulators of cognitive deficit through combined use of neuropsychological test performance and event-related potentials. *Journal of Clinical and Experimental Neuropsychology, 21*(6), 866-879.

Fabiani, M., Gratton, G., & Coles, M. G. H. (2000). Event-related brain potentials: Methods, theory, and applications. In J.T. Cacioppo, L.G. Tassinary, & G.G. Berntson (Eds.), *Handbook of Psychophysiology* (2nd ed). Cambridge University Press.

Farwell, L. A., & Donchin, E. (1991). The truth will out: Interrogative polygraphy ("lie detection") with event-related brain potentials. *Psychophysiology, 28*(5), 531-547.

Folestein, M. F., Folestein, S. E., & McHugh, P. R. (1975). "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189-198.

Friedman, D. (2000). Event-related brain potential investigations of memory and aging. *Biological Psychology, 54*(1-3), 175-206.

Greve, K. W., Bianchini, K. J., Mathias, C. W., Houston, R. J., & Crouch, J. A. (2003). Detecting malingered performance on the Wechsler Adult Intelligence Scale: Validation of Mittenberg's approach in traumatic brain injury. *Archives of Clinical Neuropsychology, 18*, 245-260.

Gronwall, D. M. A. (1977). Paced Auditory Serial-Addition Task: A measure of recovery from concussion. *Perceptual and Motor Skills, 44*, 367-373.

Guilmette, T. J., Hart, K. J., & Giuliano, A. J. (1993). Malingering detection: The use of forced-choice method in identifying organic versus simulated memory impairment. *The Clinical Neuropsychologist, 8*, 59-69.

Heaton, R. K., Smith, H. H., Lehman, R. A., & Vogt, A. T. (1978). Prospects for faking believable deficits on neuropsychological testing. *Journal of Consulting and Clinical Psychology, 46*, 892-900.

Hiscock, M., & Hiscock, C. K. (1989). Refining the forced-choice method for the detection of malingering. *Journal of Clinical and Experimental Neuropsychology, 11*, 967-974.

Inman, T. H., & Berry, D. T. R. (2002). Cross-validation of indicators of malingering: A comparison of nine neuropsychological tests, four tests of malingering, and behavioral observations. *Archives of Clinical Neuropsychology, 17*, 1-23.

Inman, T. H., Vickery, C. D., Berry, D. T. R., Lamb, D. G., Edwards, C. L., & Smith, G. T. (1998). Development and initial validation of a new procedure for evaluating adequacy of effort given during neuropsychological testing: The Letter Memory Test. *Psychological Assessment, 10*, 128-139.

Johnson, R., Jr. (1986). A triarchic model of P300 amplitude. *Psychophysiology, 23*, 367-384.

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology, 38*(3), 557-577.

Lucas, J.A. (1998). Traumatic brain injury and postconcussive syndrome. In P. J. Snyder and P. D. Nussbaum (Eds.), *Clinical Neuropsychology: A Pocket Handbook for Assessment* (pp. 243-265). Washington, D.C.: American Psychological Association

McGuire, B. E. & Shores, E. A. (1998). Malingering of memory impairment on the Colorado Priming Test. *British Journal of Clinical Psychology, 37*, 99-102.

Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology, 24*(8), 1094-1102.

Osimani, A., Alon, A., Berger, A., & Abarband, J. M. (1997). Use of the Stroop phenomenon as a diagnostic tool for malingering. *Journal of Neurology, Neurosurgery, & Psychiatry, 62*, 617-621.

Paller, K. A. (2000). Neural measures of conscious and unconscious memory. *Behavioural Neurology, 12*(3), 127-141.

Paller, K. A. (2001). Neurocognitive foundations of human memory. In D. L. Medin's (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 40, pp. 121-145). San Diego, CA: US, Academic Press.

Paller, K. A., & Kutas, M. (1992). Brain potentials during memory retrieval provide neurophysiological support for the distinction between conscious recollection and priming. *Journal of Cognitive Neuroscience, 4*, 375-391.

Pankrantz, L., & Binder, L. M. (1997). Malingering on intellectual and neuropsychological measures. In Rogers (Ed.) *Clinical assessment of malingering and deception* (2nd ed.). (pp. 223-236). New York, NY: The Guilford Press.

Picton, T. W., Bentin S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Miller, G. A., et al. (2000). Guidelines for using human event-related potentials to study

cognition: Recording standards and publication criteria. *Psychophysiology, 37*, 127-152.

Potter, D. D., Pickles, C. D., Roberts, R. C., & Rugg, M. D. (1992). The effects of scopolamine on event-related potentials in a continuous recognition memory task. *Psychophysiology, 29*, 29-37.

Rees, L. M., Tombaugh, T. N., Gansler, D. A., & Moczynski, N. P. (1998). Five validation experiments of the Test of Memory Malingering (TOMM). *Psychological Assessment, 10*(1), 10-20.

Rogers, R. (1997). *Clinical assessment of malingering and deception* (2$^{nd}$ ed.). New York: Guilford Press.

Rose, F. E., Hall, S., & Szalda-Petree, A. D. (1995). Portland Digit Recognition Test – Computerized: Measuring response latency improves the detection of malingering. *The Clinical Neuropsychologist, 9*(2), 124-134.

Rosenfeld, J. P., Angell, A., Johnson, M., & Qin, J. (1991). An ERP-based, control-question lie detector analog: Algorithms for discriminating effects within individuals' average waveforms. *Psychophysiology, 38*, 319-335.

Rosenfeld, J. P., & Ellwanger, J. (1999). Cognitive psychophysiology in detection of malingered cognitive deficit. In J. J. Sweet's, *Forensic Neuropsychology: Fundamentals and Practice* (pp. 287-311). Lisse, Netherlands: Swets & Zeitlinger Publishers.

Rosenfeld, J. P., Ellwanger, J. W., Nolan, K., Wu, S., Bermann, R. G., & Sweet, J. (1999). P300 scalp amplitude distribution as an index of deception in a simulated cognitive deficit model. *International Journal of Psychophysiology, 33*, 3-19.

Rosenfeld, J. P., Ellwanger, J., & Sweet, J. (1995). Detecting simulated amnesia with event-related brain potentials. *International Journal of Psychophysiology, 19*, 1-11.

Rosenfeld, J. P., Reinhart, A. M., Bhatt, M., Ellwanger, J., Gora, K., Sekera, M., & Sweet., J. J. (1998). P300 correlates of simulated amnesia on a matching-to-sample task: Topographic analyses of deception vs. truth-telling responses. *International Journal of Psychophysiology, 28*, 233-248.

Rosenfeld, J. P., Sweet, J. J., Chuang, J., Ellwanger, J., & Song, L. (1996). Detection of simulated malingering using forced choice recognition enhanced with event-related potential recording. *The Clinical Neuropsychologist, 10*(2), 163-179.

Rugg, M. D., & Nagy, M. E. (1989). Event-related potentials and recognition for words. *Electroencephalography and Clinical Neurophysiology, 72*, 395-406.

Slick, D. J., Sherman, E. M. S., & Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist, 13*(4), 545-561.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology and Human Learning, 6*, 174-215.

Snow, J. H., & Hooper S. R. (1994). *Pediatric traumatic brain injury*. Thousand Oaks, CL: Sage.

Strauss, E., Spellacy, F., Hunter, M., & Berry, T. (1994). Assessing believable deficits on measures of attention and information processing. *Archives of Clinical Neuropsychology, 9*, 483-490.

Tardif, H. P., Barry, R. J., Fox, A. M., & Johnstone, S. J. (2000). Detection of feigned recognition memory impairment using the old new effect of the event-related

potential. *International Journal of Psychophysiology, 36*, 1-9.

Tardif, H.P., Barry, R.J., & Johnstone, S.J. (2002). Event-related potentials reveal processing differences in honest vs. malingered memory performance. *International Journal of Psychophysiology, 46*, 147-158.

Teasdale, G., & Jennett, B. (1974). Assessment of coma and impaired consciousness: A practical scale. *Lancet, July*, 81-84.

Tombaugh, T. N. (1996). *The Test of Mental Malingering*. New York: MHS.

Tombaugh, Tombaugh, T. N. (1997). The Test of Mental Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment, 9*, 260-268.

Van Gorp, W. G., Humphrey, L. A., Kalechstein, A., Brumm, V. L., McMullen, W. J., Stoddard, M., & Pachana, N. A. (1999). How well do standard clinical neuropsychological tests identify malingering? A preliminary analysis. *Journal of Clinical and Experimental Neuropsychology, 21*(2), 245-250.

Vickery, C. D., Berry, D. T. R., Dearth, C. S., Vagnini, V. L., Baser, R. E., Cragar, D. E., & Orey, S. A. (2004). Head injury and the ability to feign neuropsychological deficits. *Archives of Clinical Neuropsychology, 599*, 1-12.

Vickery, C. D., Berry, D. T. R., Inman, T. H., Harris, M. J., & Orey, S. A. (2001). Detection of inadequate effort on neuropsychological testing: A meta-analytic review of selected procedures. *Archives of Clinical Neuropsychology, 16*, 45-73.

Vuilleumier, P, Schwartz, S., Duhoux, S., Dolan, R. J., & Driver, J. (2005). Selective Attention Modulates Neural Substrates of Repetition Priming and "Implicit" Visual Memory: Suppressions and Enhancements Revealed by fMRI. *Journal of Cognitive Neuroscience, 17*(8), 1245-1260.

Wagner, A. D., & Koutstaal, W. (2002). Priming. *Encyclopedia of the Human Brain, 4*, 27-46.

Warrington, E. K. (1984). *Recognition Memory Test Manual*. NFER-Nelson, Windsor Berkshire.

Wegesin, D. J., & Nelson, C. A. (2000). Effects of inter-item lag on recognition memory in seizure patients preceding temporal lobe resection: Evidence from event-related potentials. *International-Journal-of-Psychophysiology, 37*(3), 243-255.

Wiggins, E. C. & Brandt, J. (1988). The detection of simulated amnesia. *Law and Human Behavior, 12*(1), 57-78.

# VITA

## VICTORIA L. VAGNINI, M.S.

### BIOGRAPHICAL INFORMATION

Date of Birth: March 12, 1976
Place of Birth: Newington, CT

### EDUCATION

M.S. Clinical Psychology, University of Kentucky, Lexington, Kentucky, December 2003

M.S. Counseling Psychology, University of Kentucky, Lexington, Kentucky, May 2000

B.S. Psychology, Centre College, Danville, Kentucky, May 1998

### PUBLICATIONS

**Vagnini, V.L**., Sollman, M.J., Berry, D.T.R., Granacher, R.P., et al.  (2006).  Known-groups cross-validation of the Letter Memory Test in a compensation-seeking mixed neurologic sample.  *The Clinical Neuropsychologist, 20*, 289-304.

Jiang, Y., **Vagnini, V.**, Clark, J., & Zhang, Q. (2007). Reduced Sensitivity of Older Adults to Affective Mismatches.  *The Scientific World Journal, 7*, xxx-xxx.

Dearth, C.M.S, Berry, D.T.R., Vickery, C.D., **Vagnini, V.L.**, Baser, R., Orey, S.A., & Crager, D.E. (2005). Detection of feigned head injury symptoms on the MMPI-2 in head injured patients and community controls. *Archives of Clinical Neuropsychology, 20*(1), 95-110.

Vickery, C.D., Berry, D.T.R., Dearth, C.M.S., **Vagnini, V.L.**, Baser, R. E., Crager, D.E., & Orey, S.A.  (2004).  Head injury and the ability to feign neuropsychological deficits.  *Archives of Clinical Neuropsychology, 19*(1), 37-48.

### POSTERS

**Vagnini, V.L.**, Zhang, Q., & Jiang, Y. (2005).  *Age-related differences in affective priming*.  Poster presented at the 113th Annual Meeting of the American Psychological Association, Washington, D.C.

**Vagnini, V.L.**, Lawson, A., Liu, X, Lianekhammy, J., Bylica, K.E., Joseph, J., Kelly, T.H., & Jiang, Y.  (2005). *The effect of sensation seeking on a novelty detection task: An ERP study*.  Poster presented at the Center for Drug and Alcohol Research Translation (C-DART) meeting, Florida.

**Vagnini, V.L.**, Sollman, M.J., Berry, D.T.R., Granacher, R.P., Clark, J.A., Burton, R., & Wetter, M.W. (2005). *Known-groups cross-validation of the Letter Memory Test in a compensation-seeking mixed neurological sample*. Presented at 33rd Annual International Neuropsychological Society Meeting, Feb. 4th, St. Louis, MO.

**Vagnini, V.L.**, Sollman, M.J., Berry, D.T.R., Granacher, R.P., & Hopkins, J.N.  (2004). *Effectiveness of MMPI-2 validity scales for detection of malingered neurocognitive deficit*.  Poster presented at the 32nd Annual Meeting of the International Neuropsychological Society, Baltimore.

Jiang, Y., Lawson, A., Guo, C., **Vagnini, V.**, Clark, J., Powell, D., Andersen, A. (2006). Frontal Interaction Between Repetition Effect And Working Memory. NeuroImage, 18.

Lawson, A.L., Liu, X., Joseph, J., Kelly, T.H., **Vagnini, V.L.**, & Jiang, Y (2006). *fMRI and ERP correlates of high and low sensation seekers: Implications for drug abuse prevention*. Poster presented at the 114th Annual Meeting of the American Psychological Association, New Orleans, LA.

Jiang, Y., **Vagnini, V.**, Clark, J., & Lawson, A. (2006). *Age-related changes in brain potentials associated with old/new and repetition effects*. 11th Cognitive Aging Conference 2006 Abs, 239.

## HONORS RECEIVED

American Academy of Forensic Psychology Dissertation Award July 2006

University of Kentucky academic fellowship for 2005-2006 academic year

Behavioral Sciences internal grant for pilot study examining using an implicit memory task to detect malingered neurocognitive deficit (February 2005; February 2006)

## RESEARCH EXPERIENCE

**Research Assistant in Aging, Brain, & Cognition Lab**, Department of Behavioral Science – University of Kentucky 2004-2006

**Thesis research project, Forensic Psychiatry Institute**, Lexington, KY 2002-2004

**Research Assistant for Patient Physician Communication Study**-Department of Behavioral Sciences-University of Kentucky 1999-2000

**Research Assistant for ADHD Study Group**-Department of Psychiatry-University of Kentucky 1998-1999

## CLINICAL EXPERIENCE

**Therapist–Psychological Services Center**–Dept. of Psychology–University of Kentucky 2001-2004

**Co-leader of Dialectical Behavior Therapy (DBT) group**- Psychological Services Center 2002-2003

**Psychology Student Affiliate - Eastern State Psychiatric Hospital**, Lexington, KY 2001-2002

**Psychology Intern, Cardinal Hill Rehabilitation Hospital-Head Injury Unit**, Lexington, KY 2000-2002

**Assessment Technician, Allen Psychological Services Center** – Private practice neuropsychologist 2000-2001

## TEACHING EXPERIENCE

**Instructor for Brain and Behavior course, Department of Psychology** – University of Kentucky 2002-2003

**Teaching Assistant for graduate interviewing class**, Department of Psychology – University of Kentucky 2003-2005

**Teaching Assistant for lab courses**, Department of Psychology - University of Kentucky 2000-2007