Open Access Master's Theses

2017

The Effect of Coaching on the Ability to Identify and Pass a Measure of Insufficient Effort

Shannon Kovach
University of Rhode Island, shannonk59@gmail.com

THE EFFECT OF COACHING ON THE ABILITY TO IDENTIFY AND PASS A

MEASURE OF INSUFFICIENT EFFORT

BY

SHANNON KOVACH

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

IN

PSYCHOLOGY

UNIVERSITY OF RHODE ISLAND

2017

MASTER OF ARTS THESIS

OF

SHANNON KOVACH

APPROVED:

Thesis Committee:

Major Professor   David Faust

Grant Willis

Leslie Mahler

Nasser H. Zawia
DEAN OF THE GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND
2017

ABSTRACT

Malingering, the feigning or exaggeration of illness or injury in order to obtain an external reward, is a well-studied topic in forensic psychology and neuropsychology. In recent years, considerable research has examined coached malingering, or ways in which individuals may gather information to develop a strategy to avoid being detected as a malingerer. One example of coached malingering is using the Internet to seek out information on which measures in a typical neuropsychological battery may be used to identify malingerers. Obtaining this information may allow individuals to devise a strategy for how to perform on such tests to avoid detection. Previous studies on coached malingering often used vague or unhelpful instructions for the participants who were assigned to the coaching condition. Although these instructions may be helpful for a select subset of participants, they often pale in comparison to the amount of information on malingering measures available through the Internet.

The current study sought to examine the effect of using more specific coaching instructions, similar to the information one could obtain from a brief Internet search. There were four conditions in this study: 1) control, 2) no coaching, 3) non-specific coaching, and 4) specific coaching. Participants in all conditions except for the controls were given a hypothetical car accident scenario instructing them to malinger head injury. Participants in the non-specific and specific coaching conditions were given additional instructions warning about the use of measures to detect faking. Participants were then administered three neuropsychological tests (HVLT-R, TOMM, BVMT-R) followed by a brief questionnaire. There were three main hypotheses in this

study: 1) participants in the specific coaching condition will demonstrate two performance characteristics needed to malinger successfully more often than those who do not receive specific coaching: a) they will perform poorly on at least one of the standard tests, and b) will pass the malingering test; 2) participants in the specific coaching condition will be more accurate in identifying the purpose of the administered tests; and 3) participants in the no coaching condition will have poorer scores across all measures compared to the specific and non-specific coaching groups.

The results of the study showed that participants in the specific coaching condition had higher rates of passing the malingering measure (100%) compared to the no coaching (0%) and non-specific coaching (22.2%) conditions. In addition, all participants in the no coaching and specific coaching conditions and all but one in the non-specific coaching condition performed poorly on one or both of the genuine memory measures (HVLT-R and BVMT-R). Participants across conditions had varying rates of success in identifying the purpose of each measure. The current study indicates that individuals who are provided with information similar to that which could be obtained through a brief Internet search on a common measure of malingering can develop a strategy to pass that measure while still performing poorly on standard tests in order to successfully feign cognitive impairment.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Problematic cooperation with psychological assessment and its various manifestations, in particular malingering, have been researched extensively in neuropsychology and forensic psychology. Despite the numerous studies and publications that describe ways to detect insufficient effort, malingering continues to be a significant concern, particularly within the legal system. A plaintiff's verdict in a civil suit can result in large monetary rewards, so it is not surprising that some litigants feign or exaggerate symptoms. The potential results of error in either direction, such as depriving an injured person of just compensation due to misidentification of malingering, speaks to the seriousness of the issue. Many effort tests only seek to identify clearly deficient effort, and thus having valid tests that measure good effort would go a long way towards reducing false positive error, or erroneously identifying someone with genuine deficits as a malingerer. Despite the importance of accurately identifying the presence or absence of malingering, clear and sufficient solutions are still lacking in many circumstances (Faust, Ahern, Bridges, & Yonce, 2012).

Obtaining information on typical measures used in neuropsychological evaluation to detect falsification might help malingerers escape identification, especially if they access information on the underlying detection strategies. Malingerers might also obtain information about symptoms they should feign or exaggerate to make their claim more believable, and additionally, how these

symptoms would manifest in test performance. Such information may be available through the Internet, or as some literature suggests, even conveyed by legal counsel. The term *coached malingering* can be used broadly to convey exposure to or the provision of information that may facilitate efforts to overcome malingering detection methods.

Although many studies have been conducted on coached malingering, few arguably use sufficient or realistic coaching instructions. Often, study participants are given directions that are too vague to be helpful, or are coached on symptoms they should feign without being informed about how these symptoms might translate to test performance. More importantly, the coaching instructions in many studies are far less useful than information a person might find during a web search on how to avoid malingering detection. This limitation signifies an important gap in current research as studies on coached malingering should relate more closely to real-life situations and the steps and strategies individuals may use, especially those with greater than usual determination or resourcefulness. Further increments in the ability of professionals to differentiate genuine and malingered presentations depend heavily on enhancing proficiency evaluating these more challenging individuals.

CHAPTER 2

REVIEW OF LITERATURE

Malingering is defined in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) as "the intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives" (pp.726-727). Malingering may be especially likely to occur in legal or forensic settings. For example, in a civil case, an individual may feign cognitive impairment to gain financial compensation, or in a criminal case may attempt to feign a psychological disorder to avoid a guilty verdict. These two diverse arenas, civil and criminal law, often involve contrasting issues and concerns, and for the purposes of this thesis and based on personal preference, the focus here will be primarily on civil cases.

Considering that awards in civil suits may be large or even astronomical, it is not surprising that individuals may intentionally feign or exaggerate their symptoms to try to maximize compensation (Faust et al., 2012). Chafetz and Underhill (2013) estimated the costs associated with feigning of adult mental disorder in civil cases in 2011 to be about 20 billion dollars. Although merely an estimate, even should the figure be off by a considerable margin, it suggests that large monetary stakes are at issue. With the potential financial burden of malingering being so great, many researchers have undertaken work in this area.

One complication with determining malingering status or estimating the scope and cost of the problem is that many people who are pursuing an injury claim have genuine cognitive deficits or emotional disorders. However, if they are falsely attributing the cause of their legitimate disorders to receive financial compensation, for example, blaming a minor work accident for a disorder that was really caused by a different event, they still arguably fall within the scope of malingering. Malingering seemingly encompasses both intentionality as well as falsification (Faust et al., 2012). Among neuropsychologists, there is no standard or common set of practices for assessing malingering. One survey of practices found that 56% of neuropsychologists reported often or always using an effort measure during neuropsychological evaluations (Sharland & Gfeller, 2007). Although this is a majority, it also follows that 44% of respondents did not report routinely using any measure to detect malingering. Currently, there is no standard protocol for assessing malingering, which creates marked inconsistency in the type and validity of particular measures selected to detect malingering. Whether to perform effort testing and which measures to use is almost always left to the neuropsychologist's discretion (Heilbronner et al., 2009; Sharland & Gfeller, 2007; Slick, Tan, Strauss, & Hultsch, 2004).

Standard Tests with Embedded Indicators. One method to assess effort during a neuropsychological evaluation is to use measures or indicators that have been added to or derived from standard tests. Some of these measures are commonly included in neuropsychological batteries, such as the Minnesota Multiphasic Personality Inventory, second edition (MMPI-II; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989), the California Verbal Learning Test, second edition (CVLT-II;

Delis, Kramer, Kaplan, & Ober, 2000), and the Wechsler Adult Intelligence Scale,

fourth edition (WAIS-IV; Wechsler, 2008). These special scales or embedded indices

are typically based on performance characteristics that differ between people with

genuine cognitive impairments and suspected malingerers (Heilbronner et al., 2009).

For example, one such indicator, Reliable Digit Span (RDS), was created by summing

the longest string of digits accurately repeated over two trials for both the forward and

backward conditions on the Digit Span task of the WAIS-R (Greiffenstein, Baker, &

Gola, 1994). The commonly suggested cut-off score may be helpful in identifying

insufficient effort as it falls below the level usually observed among individuals with

genuine cognitive impairment (Mathias, Greve, Bianchini, Houston, & Crouch, 2002;

Schroeder, Twumasi-Ankrah, Baade, & Marshall, 2012).

Neuropsychologists often use these indices in combination with other measures

of effort to determine whether an examinee has cooperated sufficiently to obtain

useable results (Heilbronner et al., 2009). However, many indicators of malingering,

such as embedded measures, can lead to a relatively high rate of false positive error

(especially when multiple such indicators are used). Therefore, these indicators must

be interpreted with a high degree of caution and allow for some degree of failure

(Berthelson, Mulchan, Odland, Miller, & Mittenberg, 2013).

Stand-Alone Measures of Malingering. To date, stand-alone tests designed

specifically to test effort have often shown the highest accuracy rates for detecting

malingering (Heilbronner et al., 2009). A common strategy in test development is to

design items that seem complex, or more difficult than they are, when they are in fact

quite easy, even for individuals with neuropsychological or emotional disorders. Thus,

short of a major psychological disorder or extreme cognitive dysfunction (the cause or presence of which is often not a matter of dispute), if a person performs poorly on these measures, they are typically identified as providing insufficient effort (Bender & Rogers, 2004; Green, Rohling, Lees-Haley, & Allen, 2001; Haber & Fichtenberg, 2006). Another strategy is to evaluate the magnitude of errors, or to analyze which items a suspected malingerer has failed. The underlying rationale is that people who are intentionally feigning do not have sufficient awareness of which items to fail and which to pass. If they are failing easy items and passing hard or harder ones with unusual regularity, or if their responses to similar types of questions are inconsistent, it suggests possible malingering (Bender & Rogers, 2004).

Whatever the strategy used in developing a malingering measure, there is no one test, or combination of tests, that allows for certainty in determining malingering status. Among stand-alone measures, there exists variation in study outcomes: some measures have (or seemingly have) high levels of sensitivity and specificity, but others do not (Heilbronner et al., 2009). However, of note, many of these measures are developed using extreme groups, for example, only recruiting participants who fell neatly into either "not malingering" or "probable/definite malingering" categories. These measures are often not validated on more difficult or ambiguous cases, and thus the levels of sensitivity and specificity obtained in many studies may not generalize well to real-world applications. As may be obvious, it is rarely the clear or extreme cases that pose the greatest challenges, but rather the cases that are less extreme or more ambiguous. Thus, the high success rates in less difficult cases are unlikely to apply to the cases in which the most help is needed.

Even assuming an effort measure with strong sensitivity and specificity, it could still fail to identify many instances of malingering due to coaching. As noted, coached malingering describes the practice of being provided with, or seeking out, information on how to thwart methods to detect malingering, such as embedded or specialized tests. Coaching can take many forms. It could be as simple as a web search on symptoms of traumatic brain injury and range to such actions as obtaining information on specific tests in a neuropsychological battery in order to learn how to convincingly underperform without detection.

Types of Coaching. According to literature, attorneys may indirectly or directly instruct litigants about which measures they might or will be given and how they should perform (Brennan et al., 2009). Although such actions may seem nefarious, they may stem from an attorney's concern about false-positive errors or the dismissal of all legitimate elements of injury claims if one or another aspect of the client's presentation is exaggerated. Attorneys, such as those specializing in brain injury cases, may read extensively on the topic or attend conferences in which mental health professionals present on malingering detection methods.

Alternatively, litigants may perform their own background research and obtain access to information such as the names of malingering measures, patterns of test performance for the injury or illness they are feigning, and much more (Bauer & McCaffrey, 2006; Ruiz, Drake, Glass, Marcotte, & Van Gorp, 2002). Previously, someone attempting to feign a disorder or injury might have had to spend considerable time and effort to gather useful information on the tests they would be given in a typical neuropsychological battery. Nowadays, a brief web search on one's smart

phone or laptop could provide considerable pertinent information in a matter of minutes. Bauer and McCaffrey's (2006) study found that 44% of websites on the Test of Memory Malingering (TOMM; Tombaugh, 1996) provided sensitive information about the test, including its use for testing suboptimal effort. In addition, 12% of those websites provided links to articles that provided the cut-off score for insufficient effort and indicated that performance below that score is not typically attributable to any type of cognitive impairment (Bauer & McCaffrey, 2006).

Having access to specific information about the purpose of malingering measures and descriptions of how to avoid detection creates a critical problem for neuropsychologists. Now the question is not just how to detect malingering, but how to detect malingering in cases in which someone has been coached on avoiding detection. Considerable research has addressed this question, much of which has examined whether current malingering measures are resilient to coaching. Results have been mixed, with some studies suggesting that some measures retain high detection rates, but with other studies suggesting reduced or compromised accuracy (Dunn, Shear, Howe, & Douglas Ris, 2003; Jelicic, Ceunen, Peters, & Merckelbach, 2011; Powell, Gfeller, Hendricks, & Sharland, 2004; Rose, Hall, Szalda-Petree, & Bach, 1998; Suhr & Gunstad, 2000).

Symptom and Test Coaching. One explanation for the difference in findings across studies relates to the variability in coaching instructions provided to participants. In some studies, participants are informed of the disorder they are supposed to feign and given a list of plausible symptoms associated with the disorder (Erdal, 2004; Jelicic et al., 2011; Powell et al., 2004). This type of coaching often appears to be performed

proficiently; participants are provided with accurate information about the disorder and the associated symptoms.

In contrast, participants who are coached on beating tests are often given vague instructions or merely warned that effort tests will be administered (Erdal, 2004; Suhr & Gunstad 2000). Some studies that are classified as test-coached studies simply tell participants how to obtain a low or impaired score (i.e. fail at least half of the items) or warn participants not to exaggerate too much, but give little or no specific instruction on how to identify or beat the malingering measure (Powell et al., 2004; Rose et al., 1998). In one such study, Dunn et al. (2003) gave participants helpful instructions on how to perform convincingly on malingering measures, but did not provide specific information (such as a cut-off score or number of errors needed) for test performance to be categorized as indicating insufficient effort.

Problematic Coaching Designs. In Jelicic et al.'s (2011) study, participants in the symptom/test-coached condition were instructed to perform as though they had suffered a mild head injury and were exaggerating their symptoms. They were warned that one or more tests would be used to detect insufficient effort; one group of participants was further coached on how to perform to avoid being detected as faking. However, only two tests were given and both were malingering measures. Both measures maintained relatively high accuracy rates for the symptom/test-coached condition (80% and 87%), but no standard tests were given and participants may have felt they needed to select one or the other of the two tests on which to do poorly in order to simulate cognitive impairment. An improved design for this study might

include adding standard tests to evaluate participants' ability to identify the measures of effort.

Powell et al.'s (2004) study also had potentially serious methodological issues. Participants were given test coaching instructions specific to the Validity Indicator Profile (VIP; Frederick, 2003), but then were assessed using the TOMM, which employs a completely different strategy for malingering detection. The results showed that 96% of the test-coached participants were correctly classified as malingerers, which gives the impression that the TOMM was largely resistant to coaching efforts. Jelicic et al.'s (2011) and Powell et al.'s (2004) studies provided questionably effective coaching strategies that may not reflect the actual methods individuals use to feign cognitive deficits and attempt to avoid detection.

Although many coaching studies find that the measures of malingering under consideration are resistant to coaching, the coaching provided often seems less than sufficient when compared to what some feigners might learn or uncover. Considering that many individuals could find more detailed or accurate information on how to perform on or circumvent these measures with a quick web search, existing coaching studies may have restricted real-world applicability. Even studies providing more detailed coaching strategies (e.g. Dunn et al., 2003) may only approximate the information that rather naïve malingerers gather, but fall short of the wealth of information available on the Internet (Bauer & McCaffrey, 2006; Ruiz et al., 2002). One might consider that a litigant may have an enormous financial incentive, perhaps a million dollars or more, to feign successfully, and that even individuals with genuine injuries might embellish for any of a variety of reasons, such as the belief that they

will not be compensated justly if they play it straight. Hence, at least some individuals who contemplate exaggerating or malingering deficits might work very hard to prepare effectively.

Tan, Slick, Strauss, and Hultsch's (2002) study provided participants with a scenario describing a car accident in which they sustained a mild head injury but had subsequently recovered; the coached group was given additional instructions to malinger cognitive deficits on testing. Participants were then allowed a week prior to testing to prepare using any resources other than the experimenter or other study participants. Although over 75% of the malingering group reported using outside resources to prepare, the malingering measures still had an 80-100% detection rate. Despite being allowed to use the Internet, only 36% of the malingering group reported using this resource, and about 78-88% of those participants were still detected on malingering measures. It is doubtful that most participants obtained coaching equivalent to what at least some individuals might obtain, particularly given more time and greater incentive to malinger successfully.

Given the popularity and accessibility of the Internet, it is important to design studies using specific coaching strategies that can be obtained from a web search, especially because research on this topic is limited. The current study seeks to address this gap in the literature by providing limited, yet informative coaching to participants, such as that which can be found on the Internet, to determine the potential impact on the efficacy of malingering tests.

This study tested the following hypotheses regarding informative and realistic coaching strategies: 1) Participants who receive specific coaching will demonstrate

two performance conditions needed to malinger successfully more often than those who do not receive specific coaching: a) they will perform poorly (i.e., more than one standard deviation below the mean, a commonly used point for indicating below average performance) on at least one of the standard tests, and b) will pass (perform above the cut-offs) on the malingering test; 2) Participants who receive specific coaching instructions versus those who do not will demonstrate greater accuracy identifying the purpose of the administered tests (e.g. effort test vs. genuine memory test) compared to each of the other groups; and 3) Participants in the no coaching condition will have poorer scores across all measures compared to the specific and non-specific coaching groups.

CHAPTER 3

METHODOLOGY

Participants. The study was advertised to undergraduate students in introductory

psychology courses at the University of Rhode Island; students in these introductory

courses are encouraged to participate in research for extra credit. Forty-one

participants from undergraduate psychology classes volunteered to participate in the

study. Two participants in the non-specific coaching condition were initially excluded

from the analyses due to low compliance or confusion about task demands, determined

by their ratings on the last few questionnaire items, which assessed their understanding

of the condition they were in and their subjective perception of how successful they

were at complying with given instructions. An additional two participants (one in the

control condition and one in the specific coaching condition) were identified as

extreme outliers in the dataset during preliminary analyses. The primary researcher

reviewed their files for data-entry related errors and noticed that the behavioral

observations completed by the test administrator revealed that both participants had

low interest/engagement in the task and were likely non-compliant with the given

instructions; thus these outliers were also removed from the dataset.

The final sample size included in the analyses was 37. Most participants ($n =$

34) were between the ages of 18 and 24, with 3 participants 25 or older. The majority

of participants identified as white ($n = 28$), with 4 identifying as multiracial, 3 as

other, 1 as black, and 1 as Hispanic/Latino. Thirty-one participants identified as female and six identified as male.

Procedures. The current study examined whether minimal but informative coaching instructions allowed participants to avoid detection on a malingering measure while performing poorly on standard neuropsychological tests, two basic conditions that typically must be met to successfully feign cognitive impairment. This was an experimental study with one independent variable, coaching, with four levels: 1) Control, 2) No Coaching, 3) Non-Specific Coaching, and 4) Specific Coaching. The control group received instructions to provide their best effort on testing. The other three experimental groups received a hypothetical car accident scenario, which stated that the participant was involved in a car accident and should feign traumatic brain injury while completing the tests in order to maximize financial gain from a lawsuit (see Appendix A).

The no coaching condition was provided no further instruction. The non-specific coaching condition was given additional information that included a warning that there will be a measure to detect faking and basic instructions on how to avoid detection. The specific coaching condition also warned participants that there will be a measure to detect faking, but the participants received a more specific description of this measure and how they should perform in order to avoid detection. Participants were asked whether or not they understood what they were being asked to do after reading through the instructions. If they needed further information or clarification, test administrators provided additional information from a set script (see Appendix B). They were then asked again if they understood the task instructions; if they still did

not understand the instructions, they were still allowed to participate in the study but their data were only used for exploratory analysis. Participants were randomly assigned to testing conditions prior to the start of testing.

Participants were tested individually in a quiet room on campus (at the Psychological Consultation Center), where they could work without disruption. At the start of the study, participants were given a consent form explaining their rights (e.g., their right to refuse to participate or to leave at any point); testing began after participants signed the consent forms. Consent forms were stored separately from the participants' test data and questionnaires to ensure anonymity.

Participants were administered three different tests commonly found in neuropsychology testing batteries: the Hopkins Verbal Learning Test - Revised (HVLT-R), the Test of Memory Malingering (TOMM), and the Brief Visuospatial Memory Test - Revised (BVMT-R). After participants finished the three tests, they completed a follow-up questionnaire. This questionnaire began with a few items to collect non-identifiable demographic information (age range, race/ethnicity, and gender). It then asked participants to select the primary purpose of each measure they were given (either genuine memory or effort/faking), rate their confidence in their selections, and how difficult they found it to identify the measure(s) of malingering. Participants were asked to rate the relative difficulty of each measure and the clarity of the instructions for the study. Finally, participants were asked to identify from among three options the instructions they were given at the start of the study and rate how successful they thought they were at complying with the given instructions (see Appendix C).

A brief pilot study was conducted to test the clarity and potential effectiveness of the specific coaching instructions. Five graduate students were recruited, provided with the hypothetical car accident scenario followed by the written specific coaching instructions, and then asked to complete the TOMM. This targeted a single IV condition and one test to assess the potential impact of the specific coaching instructions and the participants' ability to follow the instructions and pass the TOMM. Other instructions and variables were not examined given the available background studies on these matters. Power analysis for this study was conducted after the pilot study to provide an effect size estimate (most studies on coached malingering do not report effect sizes, making effect size estimates tenuous in these circumstances).

The minimum sample size necessary to assess study hypotheses was calculated *a priori* using G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007). Effect size was calculated using TOMM Trial 2 average scores for the four participant conditions (see below for a description of the TOMM). For the coaching condition, the average score was obtained from the pilot study (48.2), and average scores for the other three conditions were calculated from the literature (control = 50, no coaching = 31.5, and non-specific coaching = 35.6). Using a power level of .80 and a calculated effect size of 2.7 (large effect sizes are common, when reported, in the general malingering literature), the estimated total sample size needed was only 8 participants. However, to err on the side of caution given the tenuousness of the assumptions and limits of the pilot study, a larger target sample of 40 participants was selected.

Measures. The Test of Memory Malingering (TOMM; Tombaugh, 1996) is a 50-item visual recognition test designed to distinguish between genuine responding and insufficient effort. The TOMM consists of two learning trials and what is described as an optional retention trial. Pictures of 50 common items are shown individually, one after another until the set is completed, followed by a series of 50 two-choice recognition items (one choice duplicating a previously shown item and the other an item not previously shown). The TOMM has demonstrated high levels of sensitivity and specificity in distinguishing malingerers from genuine responders using the cut-off score recommended in the manual on both the second learning trial and the retention trial (with the first learning trial used as an ancillary measure) (Haber & Fichtenberg, 2006; Jelicic et al., 2011; Rees, Tombaugh, Gansler, & Moczynski, 1998; Wisdom, Brown, Chen, & Collins, 2012).

Studies suggest that the TOMM is insensitive to the effects of education, age, and various types of cognitive and psychological impairments, which makes it an effective measure of effort (Haber & Fichtenberg, 2006; Rees et al., 1998). In the current study, TOMM trials 1 and 2 were administered to participants, but not the optional retention trial in an effort to reduce total administration time. Participants' score on trial 2 was used to determine if they were providing sufficient or insufficient effort using the recommended cut-off score.

The Hopkins Verbal Learning Test- Revised (HVLT-R; Brandt & Benedict, 2001) is a 12-item test of verbal learning and memory. The test consists of three free recall trials. Each correctly recalled word is given a score of 1 for a maximum score of 12 for each trial; a total score is calculated by summing the three trial scores. The

17

HVLT-R has shown good construct and concurrent validity and modest to good test-retest reliability (O'Neil-Pirozzi, Goldstein, Strangman, & Glenn, 2012; Shapiro, Benedict, Schretlen, & Brandt, 1999; Woods et al., 2005). The HVLT-R has demonstrated good specificity and sensitivity for detecting various cognitive impairments in older adults and traumatic brain injury in adults (De Jager, Schrijnemaekers, Honey, & Budge, 2009; O'Neil-Pirozzi et al., 2012; Shapiro et al., 1999). In the current study, the HVLT-R was used as a brief cognitive test for the purpose of score comparisons among the other tests; the retention trial was omitted to limit administration time.

The Brief Visuospatial Memory Test- Revised (BVMT-R; Benedict, 1997) is a measure of visual memory similar in design to the HVLT-R. Participants are shown a visual stimulus with six geometric figures for 10 seconds and then asked to draw as many of the figures as they can recall; this trial is repeated twice more. Each trial is given a score based on location and accuracy of figures for a maximum score of 12 for each trial; a total score is calculated by summing the three trial scores. The BVMT-R is highly correlated with other measures of visual memory and visuospatial construction and has shown good construct validity and test-retest reliability (.96 - .97 for the learning trials) (Benedict, Schretlen, Groninger, Dobraski, & Shpritz, 1996; Hubley & Stinnett 2007; Strauss, Sherman, Spreen, & Spreen, 2006). In the current study, the total score from the BVMT-R was used to assess visual memory, and the retention trial was omitted to limit administration time.

CHAPTER 4

RESULTS

Statistical Analyses. Preliminary analyses were conducted using SPSS version 24

(IBM Corp., 2016). Group comparisons indicated that demographic variables (i.e.,

age, race/ethnicity, sex) were similarly distributed across coaching conditions (see

Table 1). Statistical assumptions of analysis of variance (ANOVA) were also

examined, including independence of cases, normality, and homogeneity of variance.

No assumptions were violated.

Table 2 provides the mean scores for the three tests in relation to the four

conditions. The control condition had the highest scores across all three measures. The

three coaching conditions had poorer performance on the HVLT-R and BVMT-R.

While the no coaching and non-specific coaching conditions showed poor

performance on the TOMM, the specific coaching condition had scores comparable to

the control condition.

Hypotheses 1 (a and b) and 3 were tested using a series of one-way between

subjects ANOVAs to determine if scores on each measure (TOMM Trial 2 score,

HVLT-R total score, BVMT-R total score) differed significantly across coaching

conditions. Table 3 presents these results, which indicate that scores differed

significantly between coaching conditions for the HVLT-R ($F(3,33) = 12.59$, $p < .001$)

and the TOMM ($F(3,33) = 30.01$, $p < .001$), but did not quite reach significance for the

BVMT-R ($F(3,33) = 2.77$, $p = .057$). In order to determine which groups performed

significantly different from one another; Tukey post-hoc tests were examined. For the HVLT-R, the control condition was significantly different at the $p < .05$ level from the three coaching conditions; however, the three coaching conditions were not significantly different from one another (see Table 4). This indicates that participants in the control condition outperformed all other participants on the HVLT-R, while participants in the coaching conditions performed at approximately the same level.

The Tukey post-hoc analysis for the TOMM found that the specific coaching condition was significantly different from the no coaching and non-specific coaching conditions at the $p < .05$ level, but was not significantly different from the control condition (see Table 5). This suggests that participants in the specific coaching condition performed about the same as the controls and that these two groups outperformed the no coaching and non-specific conditions.

In order to determine if participants performed poorly on the genuine memory tests, the test manuals for the HVLT-R and BVMT-R were used to identify cut-off scores; any participant who scored below these cut-off scores was considered to be performing poorly. Table 6 shows the number of participants in each condition who failed/passed the TOMM and whether they performed poorly on zero, one, or both of the genuine memory tests.

The participant counts contained in this table indicate that all participants in the control condition passed the TOMM, but two out of nine performed poorly on one of the genuine memory tests. Of note, these two controls performed poorly on the BVMT-R, all controls performed normally on the HVLT-R. In the no coaching and non-specific coaching conditions, most participants failed the TOMM and also

performed poorly on one or both genuine measures of memory. All participants in the specific coaching condition passed the TOMM; in addition, six out of nine participants performed poorly on one of the memory measures and the other three participants performed poorly on both.

Using the participant counts reported in Table 6, the classification accuracy of the TOMM trial 2 recommended cut-off score was calculated. There were no false positive errors, with 100% of the controls being correctly classified as non-malingerers. For the coaching conditions, 100% of the participants in the no coaching condition were correctly classified as malingering. In the non-specific coaching condition, seven out of nine participants were correctly classified as faking and 2 avoided detection. In the specific coaching condition, all 9 participants avoided detection. The sensitivity of the TOMM Trial 2 score is 77.78% in the non-specific coaching condition and 0% in the specific coaching condition, indicating that the cut-off score was able to detect most malingers in the non-specific coaching condition, but misidentified all participants in the specific coaching condition as non-malingerers.

Eta-squared was calculated to examine effect size for the significant ANOVAs. Using Cohen's (1988) benchmarks to interpret eta-squared, results indicated that there were large effect sizes for coaching condition on both the HVLT-R ($\eta^2 = 0.53$) and the TOMM ($\eta^2 = 0.73$) Although the BVMT-R ANOVA did not quite reach significance, eta-squared was calculated and also found a large effect ($\eta^2 = 0.20$). This indicates that 20% of the variance in the BVMT-R, 53% of the variance in HVLT-R performance, and 73% of the variance in TOMM performance were accounted for by coaching condition. Given that participants in the specific coaching condition had a similar rate

of passing the TOMM as controls, (while most participants in the other two coaching groups failed the TOMM) the results suggest that participants in the specific coaching group were able to avoid being detected as faking. Concurrently, participants in the specific coaching condition performed poorly on one or both of the genuine memory measures, (similar to the no coaching and non-specific coaching groups) satisfying the second condition to be considered a successful malingerer.

For hypothesis 2, three chi-square tests of independence were conducted to determine the relationship between condition and accuracy in identifying the purpose of each measure (e.g. choosing the TOMM as the effort test). Chi-square tests of independence revealed a significant relationship between coaching condition and correctly identifying the purpose of the HVLT-R (χ^2 (3, N = 37) = 9.85, p =.02) and TOMM (χ^2 (3, N = 37) =8.75, p =.03), but not the BVMT-R (χ^2 (3, N=37) = 7.07, p = .07). Participants in the control condition were more likely to accurately identify the HVLT-R as a genuine memory test, and participants in the no coaching condition had a higher likelihood of accurately identifying the TOMM as an effort/faking test (see Table 7). Cramer's V was calculated as a measure of effect size for the association between coaching condition and correctly identifying the purpose of the HVLT-R (φc = 0.52), the TOMM (φc = 0.49), and the BVMT-R (φc = 0.44) and found large effects for all three.

CHAPTER 5

DISCUSSION

Due to the substantial financial burden and harm created by false negative and

false positive errors in malingering detection, it is critical that valid measures of effort

and symptom validity are incorporated into neuropsychological assessment.

Furthermore, given the possibility, if not the likelihood, of litigants seeking

information about malingering measures and how to avoid detection, these measures

should be resistant to coaching. Research has demonstrated that various measures of

effort, at least under certain conditions, often prove resistant to coaching. However,

the coaching instructions given to participants are often vague and do not provide a

clear strategy (Dunn et al., 2003; Jelicic et al., 2011; Powell et al., 2004; Rose et al.,

1998). In addition, the coaching instructions provided probably do not compare to the

amount of information that might be available through a brief Internet search on the

respective measure.

The results of the current study provide support for the first hypothesis –

participants who received specific coaching performed poorly on at least one of the

standard tests and passed the malingering test more often than participants who did not

receive specific coaching. All nine participants in the specific coaching condition

avoided detection on the TOMM compared with 2 out of 9 in the non-specific

coaching condition, and zero participants in the no coaching condition. This finding is

inconsistent with Powell et al.'s (2004) results, which showed that the TOMM trial 2

cut-off score was able to detect 92.6% of symptom coached and 96.0% of test coached malingerers. Similarly, the outcome of the present study is also inconsistent with Jelicic et al.'s (2011) results, which showed that 87% of the symptom coached and 80% of the symptom and test coached participants were identified using the TOMM trial 2 recommended cut-off score.

In addition to passing the TOMM, participants in the specific coaching condition performed similarly to the other two coaching conditions on the HVLT-R and BVMT-R, and all coaching conditions performed significantly worse than the control condition. This indicates that participants in the specific coaching condition did not adopt a more conservative approach to testing and perform normally across all three tests; rather, it suggests that these participants were able to identify the test of malingering and perform normally on this measure while still feigning impairment on one or both of the two genuine memory measures.

The second hypothesis, that individuals in the specific coaching condition would be better able to detect the true purpose of the tests, was not supported. Across all conditions, participants had variable success identifying the true purpose of the measures. While there were some significant relationships based on condition as discussed in the results section, it seems likely this was a result of random responding or a tendency to select the same purpose for all measures. For example, most participants in the control condition identified all the measures as genuine memory tests and most participants in the no coaching condition selected the purpose of all measures as effort/faking.

The lack of support for the second hypothesis was particularly surprising given

that participants in the specific coaching group were provided with a description of the

TOMM and told that it was a measure designed to detect faking. Furthermore, most

participants in this condition were able to identify the TOMM during testing, and their

performance was consistent with the coaching provided in the instructions (i.e. do

your best on this measure), indicating that they were complying with their given

instructions. However, seven out of nine participants in the specific coaching

condition selected the purpose of the TOMM as a measure of genuine memory. One

possible explanation is that the question might have been poorly worded and

participants were confused about what they were being asked to rate. For example,

participants could have interpreted the question to be assessing how they approached

the test, so they might have selected the purpose of a measure as effort/faking if they

were intentionally performing poorly or selected it as a genuine memory test if they

were providing their best effort. It is not clear why there was a disconnect between

participants successfully following their specific coaching instructions and performing

in a way so as to avoid detection on the TOMM, yet failing to identify the TOMM as a

measure of effort/faking on the questionnaire.

The results of the current study partially support the third and final hypothesis -

that participants in the no coaching condition would demonstrate poorer performance

across all measures compared to the other two coaching conditions. Participants in the

no coaching condition had the poorest performance on the TOMM (M =27.30) and the

BVMT-R (M =16.90), but not the HVLT-R (M =18.00); the specific coaching

condition had the poorest performance on HVLT-R (M =16.78). Even though

participants across all coaching conditions were warned not to get caught faking, without further information on how to avoid detection, participants in the no coaching condition feigned severe impairment across all measures. This condition would be similar to a naïve approach to malingering in which the person feigning impairment has no apparent strategy for feigning deficits and has not been coached, or at least successfully coached, on how to perform to avoid detection.

The results of the current study suggest that information that might be gleaned through a brief web search on the TOMM might well be sufficient to avoid being detected as malingerers (at least on that test). As mentioned in the results section, the sensitivity of the TOMM Trial 2 cut-off score is 100% for the no coaching condition and 77.78% for the non-specific coaching condition, which indicate good to ideal levels of sensitivity for a measure. However, the sensitivity drops to 0% for the specific coaching condition, which suggests that all participants in this condition were able to avoid being detected using the current cut-off score for the TOMM trial 2. At the same time, all nine participants in the specific coaching condition also performed poorly on at least one of the standard measures, indicating that all participants in this condition met both criterion needed to be considered successful malingerers. This is an important finding because it suggests that given more detailed coaching, participants are able to successfully malinger while avoiding being detected as malingerers.

The level of sensitivity for the specific coaching condition is also well below the level reported in past coaching studies using the TOMM (Jelicic et al., 2011; Powell et al., 2004). This suggests that although the TOMM has been shown to be resistant to coaching strategies in past studies, it does not retain the same level of sensitivity when

26

participants are provided with more descriptive information on how to identify the test in a battery and how to avoid detection, conditions that might be met in a fair percentage of litigated or actual cases.

There are several strengths and limitations to this study. One strength was the introduction of a novel specific coaching condition designed to be comparable to the type of information that could be obtained from an Internet search. Another strength of the current study was the follow-up questionnaire in which participants were given the chance to rate the clarity of instructions, to try to identify the condition to which they were assigned, and to rate their self-perceived success in following the task demands. This allowed the researcher to determine whether the coaching instructions were confusing or hard to follow, a potential confound in the study, as well as to assess whether participants understood which condition they were in and how well they thought they were able to follow their instructions. A third strength of the study was having research assistants who were blind to the study hypotheses administer the tests to all but one of the participants, which minimized potential bias in the administration of the tests and questionnaire. (One participant was assessed by the primary researcher due to a research assistant's last minute cancellation; however, the participant was in the control condition and their performance was likely not influenced by the test administrator knowing the study hypotheses.)

One limitation in the current study is the restricted external validity due to sample demographics. The current sample was comprised primarily of 18 to 24-year-old, white, female college students in introductory psychology classes. This sample is likely not reflective of the population of real-life malingerers. Another limitation is

that the follow-up questionnaire was designed for this study and is not based on a validated measure. To the researcher's knowledge, there are no existent, validated measures that assess the constructs the researcher was interested in gathering. Nevertheless, using a non-validated measure creates various uncertainties, including the possibility that some questionnaire items may be poorly worded and may not accurately assess the construct it is aiming to capture.

The findings of this study suggest that information parallel to that which can be obtained from the Internet can assist malingerers in developing an effective strategy to feign impairment while avoiding detection on the TOMM. Future research could extend the current findings by giving participants access to the Internet and leaving them on their own to develop their own strategies, rather than being provided with explicit coaching instructions. Utilizing this study design would help determine if a layperson can develop a successful coaching strategy using only information obtained through a web search.

Another avenue for future research would be to develop a new measure or revise current malingering measures to be more resistant to coaching. One such way to achieve this goal would be to reduce transparency by increasing the complexity of current measures. An example of this would be to include more difficult items and then randomize item difficulty. Most tests are designed to increase in difficulty as you go, which allows individuals trying to feign impairment to select a point at which to start performing poorly and miss more complex items. However, with randomized difficulty, those trying to malinger would likely have trouble trying to devise a consistent performance strategy when challenging items are interspersed with very

easy items. Having measures of effort with higher ceilings would also allow for measuring degrees of effort rather than a binary good effort and poor effort.

TABLES

Table 1

Demographics of the Sample by Coaching Condition

	Control	No Coaching	Non-Specific Coaching	Specific Coaching
Age				
18-24	9	9	9	7
25 and over	0	1	0	2
Race/Ethnicity				
White	7	7	7	7
Black	1	0	0	0
Hispanic/Latino	0	1	0	0
Multiracial	1	1	1	1
Other	0	1	1	1
Gender				
Female	7	8	7	9
Male	2	2	2	0

Note. The follow-up questionnaire included other demographic options that were not endorsed by any participant (*Asian* and *Native American* were included as options for race/ethnicity, while *Other* and *Prefer Not to Answer* were included as options for gender).

Table 2

Average Scores on HVLT-R, TOMM, and BVMT-R by Coaching Condition

	HVLT-R		TOMM		BVMT-R	
	Mean	SD	Mean	SD	Mean	SD
Control	29.67	2.06	50.00	0.00	26.22	7.09
No Coaching	18.00	5.64	27.30	8.97	16.90	7.22
Non-Specific Coaching	18.89	4.73	40.00	7.14	22.44	5.41
Specific Coaching	16.78	6.55	48.78	1.79	21.56	8.40

Note. SD= Standard Deviation.

Table 3

One-Way Analysis of Variance of HVLT-R, TOMM, and BVMT-R Scores by Coaching Condition

	SS	*df*	*MS*	*F*	*p*
HVLT-R					
Between groups	964.37	3	321.46	12.59	.000
Within groups	842.44	33	25.53		
Total	1806.81	36			
TOMM					
Between groups	3158.67	3	1052.89	30.01	.000
Within groups	1157.66	33	35.08		
Total	4316.32	36			
BVMT-R					
Between groups	419.53	3	139.84	2.77	.057
Within groups	1668.90	33	50.57		
Total	2088.43	36			

Note. SS= sum of squares; df= degrees of freedom; MS= mean square.

Table 4

Tukey HSD Comparisons of HVLT-R Scores between Coaching Conditions

				95% Confidence Interval	
Group 1	Group 2	Mean Diff.	Std. Error	Lower Limit	Upper Limit
Control	No Coaching	11.67*	2.32	5.39	17.95
	Non-Specific Coaching	10.78*	2.38	4.34	17.22
	Specific Coaching	12.89*	2.38	6.45	19.33
No Coaching	Control	-11.67*	2.32	-17.95	-5.39
	Non-Specific Coaching	-0.89	2.32	-7.17	5.39
	Specific Coaching	1.22	2.32	-5.06	7.50
Non-Specific Coaching	Control	-10.78*	2.38	-17.22	-4.34
	No Coaching	0.89	2.32	-5.39	7.17
	Specific Coaching	2.11	2.38	-4.33	8.55
Specific Coaching	Control	-12.89*	2.38	-19.33	-6.45
	No Coaching	-1.22	2.32	-7.50	5.06
	Non-Specific Coaching	-2.11	2.38	-8.55	4.33

Note. * $p < 0.05$

Table 5

Tukey HSD Comparisons of TOMM Scores between Coaching Conditions

| Group 1 | Group 2 | Mean Diff. | Std. Error | 95% Confidence Interval | |
				Lower Limit	Upper Limit
Control	No Coaching	22.70*	2.72	15.34	30.06
	Non-Specific Coaching	10.00*	2.79	2.45	17.55
	Specific Coaching	1.22	2.79	-6.33	8.77
No Coaching	Control	-22.70*	2.72	-30.06	-15.34
	Non-Specific Coaching	-12.70*	2.72	-20.06	-5.34
	Specific Coaching	-21.48*	2.72	-28.84	-14.12
Non-Specific Coaching	Control	-10.00*	2.79	-17.55	-2.45
	No Coaching	12.70*	2.72	5.34	20.06
	Specific Coaching	-8.78*	2.79	-16.33	-1.23
Specific Coaching	Control	-1.22	2.79	-8.77	6.33
	No Coaching	21.48*	2.72	14.12	28.84
	Non-Specific Coaching	8.78*	2.79	1.23	16.33

Note. * $p < 0.05$

Table 6

Participant Counts of Performance on TOMM and Genuine Memory Tests Based on Condition

Performance on Memory Tests	Control	No Coaching	Non-Specific Coaching	Specific Coaching
Fail TOMM				
Normal on both	0	0	1	0
Poor on 1	0	2	1	0
Poor on both	0	7	5	0
Pass TOMM				
Normal on both	7	0	0	0
Poor on 1	2	0	2	6
Poor on both	0	0	0	3

Table 7

Participant Counts of Correct Identification of Purpose of Measure by Coaching Condition

	HVLT-R		TOMM		BVMT-R	
	Correct	Incorrect	Correct	Incorrect	Correct	Incorrect
Control	9	0	2	7	7	2
No Coaching	3	7	8	2	3	7
Non-Specific Coaching	5	4	4	5	2	7
Specific Coaching	5	4	2	7	5	4

APPENDICES

Appendix A. Head injury scenario.

You were recently involved in a car accident. Your car was hit from behind, even though you were completely stopped at a red light. Although you were not at fault at all, the other driver didn't even apologize and put all the blame on you. While you only suffered minor injuries, like a sore neck, you remember you had a friend who pretended to have a bad concussion and got a lot of money that paid for all college costs and more. You decide that you are going to pretend to have a bad concussion that has caused a bunch of problems. After all, so many other individuals have gotten money in legal cases, it is just a big insurance company's money and they will never miss it, and there have been so many other times you have bent over backwards to be fair and have ended up on the short end of things.

You know that individuals with concussions often have serious problems with headaches, and with thinking, concentrating, and especially remembering things. You are being tested by a psychologist, who will give you tests to measure your mental abilities. When you take the tests that follow, make sure you perform poorly to show you have serious problems – the worse the better -- but not to the point that you might get caught faking the results.

* For more information on the non-specific and specific coaching instructions provided to participants, please contact the primary researcher.

Appendix B. Further Instructions

After participants read through the instructions for their given condition (with the exception of the control condition in which participants receive verbal instructions to provide their best effort on testing), the examiner will ask if they understand what they are being asked to do. If participants state that they are unsure of what they are being asked to do or if they appear to be uncertain, the examiner will provide further information (provided below) based on their assigned condition. They will then be asked again if they understand what they are being asked to do, and if they still cannot understand the instructions, they will be excused from the study.

Control Condition

Examiner: "You will be given a series of tests, do the best that you can on each measure."

No Coaching Condition:

Examiner: "In the story you were given, you were involved in a car accident. Although you had only minor injuries, you decide to pretend to have a concussion to get money. You will be given a series of tests and should perform poorly to show you have serious problems from the injury; however, you should not perform so poorly that it seems you are faking."

* For more information on the non-specific and specific coaching instructions provided to participants, please contact the primary researcher.

Appendix C. End of Study Questionnaire.
Demographics

Age
- 17 and under
- 18-24
- 25 and over

Race/Ethnicity
- White
- Black
- Hispanic/Latino
- Asian
- Native American
- Multiracial
- Other

Gender
- Female
- Male
- Other
- Prefer not to answer

Questions

What do you think the primary purpose of Test 1 (the list of words) is?
- Effort/Faking Test
- True Memory Test

How confident are you that you selected the correct primary purpose of the test?

0	10	20	30	40	50	60	70	80	90	100
Not Confident At All										Completely Confident

What do you think the primary purpose of Test 2 (the pictures of objects) is?
- Effort/Faking Test
- True Memory Test

How confident are you that you selected the correct primary purpose of the test?

0	10	20	30	40	50	60	70	80	90	100
Not Confident At All										Completely Confident

What do you think the primary purpose of Test 3 (the geometric figures) is?
- Effort/Faking Test
- True Memory Test

How confident are you that you selected the correct primary purpose of the test?

0	10	20	30	40	50	60	70	80	90	100
Not Confident At All										Completely Confident

How difficult was it to identify which test(s) were designed to detect insufficient effort (faking?)

1	2	3	4	5	6	7
Very Easy						Very Difficult

How would you rate the difficulty of the first test (list of words)?

1	2	3	4	5	6	7
Very Easy						Very Difficult

How would you rate the difficulty of the second test (pictures of objects)?

1	2	3	4	5	6	7
Very Easy						Very Difficult

How would you rate the difficulty of the third test (geometric figures)?

1	2	3	4	5	6	7
Very Easy						Very Difficult

How clear were the instructions you were given at the start of this study?

1	2	3	4	5	6	7
Very Easy						Very Difficult

What did the instructions you were given at the start of the study ask you to do?

- Provide my best effort on the tests

- Fake head injury on testing

- Fake head injury on testing and try to avoid being caught faking

How successful do you think you were at following the instructions you were given?

1	2	3	4	5	6	7
Not Successful						Very Successful

BIBLIOGRAPHY

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental

  disorders: DSM-5. Washington, D.C: American Psychiatric Association.

Bauer, L & McCaffrey, R. J. (2006). Coverage of the Test of Memory Malingering,

  Victoria Symptom Validity Test, and Word Memory Test on the Internet: Is

  test security threatened? *Archives of Clinical Neuropsychology*, *21*, 121-126.

Bender, S. D., & Rogers, R. (2004). Detection of neurocognitive feigning:

  Development of a multi-strategy assessment. *Archives of Clinical*

  Neuropsychology, 19(1), 49-60.

Benedict, R. H. (1997). *Brief Visuospatial Memory Test--Revised: Professional*

  manual. Psychological Assessment Resources.

Benedict, R. H., Schretlen, D., Groninger, L., Dobraski, M., & Shpritz, B. (1996).

  Revision of the Brief Visuospatial Memory Test: Studies of normal

  performance, reliability, and validity. *Psychological Assessment, 8*(2), 145-

  153.

Berthelson, L., Mulchan, S. S., Odland, A. P., Miller, L. J., & Mittenberg, W. (2013).

  False positive diagnosis of malingering due to the use of multiple effort tests.

  Brain Injury, 27(7-8), 909-916.

Brandt, J., & Benedict, R. H. (2001). *Hopkins Verbal Learning Test--Revised:*

  Professional manual. Psychological Assessment Resources.

Brennan, A. M., Meyer, S., David, E., Pella, R., Hill, B. D., & Gouvier, W. D. (2009).

The vulnerability to coaching across measures of effort. *The Clinical Neuropsychologist, 23*(2), 314-328.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A, & Kaemmer, B. (1989). *The Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring*. Minneapolis, MN: University of Minnesota Press.

Chafetz, M., & Underhill, J. (2013). Estimated costs of malingered disability. *Archives of Clinical Neuropsychology, 28*(7), 633-639.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, *2nd edition*. Hillsdale. NJ: Lawrence Earlbaum Associates.

De Jager, C. A., Schrijnemaekers, A. C., Honey, T. E., & Budge, M. M. (2009). Detection of MCI in the clinic: evaluation of the sensitivity and specificity of a computerised test battery, the Hopkins Verbal Learning Test and the MMSE. *Age & Ageing, 38*(4), 1-6.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *California Verbal Learning Test—2nd edition: Manual*. San Antonio, TX: The Psychological Corporation.

Dunn, T. M., Shear, P. K., Howe, S., & Douglas Ris, M. (2003). Detecting neuropsychological malingering: effects of coaching and information. *Archives of Clinical Neuropsychology, 18*, 121-134.

Erdal, K. (2004). The effects of motivation, coaching, and knowledge of neuropsychology on the simulated malingering of head injury. *Archives of Clinical Neuropsychology, 19*, 73-88.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible

statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Faust, D., Ahern, D. C., Bridges, A. J., & Yonce, L. J. (2012). *Assessment of malingering and falsification: Conceptual foundations and sources of error*. In C.R. Reynolds & A.M. Horton Jr. (Eds.), Detection of malingering during head injury litigation (1-67). New York, NY: Springer Science + Business Media.

Frederick, R. I. (2003). *Validity Indicator Profile: Manual*. Minneapolis: Pearson.

Green, P., Rohling, M. L., Lees-Haley, P. R., & Allen, L. M., III. (2001). Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury, 15*(12), 1045-1060.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment, 6*(3), 218-224.

Haber, A. H., & Fichtenberg, N. L. (2006). Replication of the Test of Memory Malingering (TOMM) in a traumatic brain injury and head trauma sample. *The Clinical Neuropsychologist, 20*(3), 524-532.

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., Millis, S. R., & Conference Participants 1. (2009). American Academy of Clinical Neuropsychology Consensus Conference Statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist, 23*(7), 1093-1129.

Hubley, A.M., Stinnett, T.A. (2007) Review of the Brief Visuospatial Memory Test –

Revised. In B. S. Plake, B. S., J. C. Impara, & R. A. Spies (Eds.), *The fifteenth mental measurements yearbook* [Electronic version]. Retrieved from the Mental Measurements Yearbook with Tests in Print online database.

IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.

Jelicic, M., Ceunen, E., Peters, M. J.V., & Merckelbach, H. (2011). Detecting coached feigning using the test of memory malingering (TOMM) and the structured inventory of malingered symptomatology (SIMS). *Journal of Clinical Psychology, 67*, 850-855.

Mathias, C. W., Greve, K. W., Bianchini, K. J., Houston, R. J., & Crouch, J. A. (2002). Detecting malingered neurocognitive dysfunction using the reliable digit span in traumatic brain injury. *Assessment, 9*(3), 301-308.

O'Neil-Pirozzi, T. M., Goldstein, R., Strangman, G. E., & Glenn, M. B. (2012). Test-re-test reliability of the Hopkins Verbal Learning Test-Revised in individuals with traumatic brain injury. *Brain Injury, 26*(12), 1425-1430.

Powell, M. R., Gfeller, J. D., Hendricks, B. L., & Sharland, M. (2004). Detecting symptom- and test-coached simulators with the test of memory malingering. *Archives of Clinical Neuropsychology, 19*, 693-702. doi:10.1016/j.acn.2004.04.001

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.

Rose, F. E., Hall, S., Szalda-Petree, A. D., & Bach, P. J. (1998). A comparison of four

tests of malingering and the effects of coaching. *Archives of Clinical Neuropsychology, 13*, 349-363.

Ruiz, M. A., Drake, E. B., Glass, A., Marcotte, D., & Van Gorp, W. G. (2002). Trying to beat the system: Misuse of the internet to assist in avoiding the detection of psychological symptom dissimulation. *Professional Psychology: Research and Practice, 33*(3), 294-299.

Schroeder, R. W., Twumasi-Ankrah, P., Baade, L. E., & Marshall, P. S. (2012). Reliable Digit Span A Systematic Review and Cross-Validation Study. *Assessment, 19*(1), 21-30.

Shapiro, A. M., Benedict, R. H., Schretlen, D., & Brandt, J. (1999). Construct and concurrent validity of the Hopkins Verbal Learning Test–revised. *The Clinical Neuropsychologist, 13*(3), 348-358.

Sharland, M. J., & Gfeller, J. D. (2007). A survey of neuropsychologists' beliefs and practices with respect to the assessment of effort. *Archives of Clinical Neuropsychology, 22*(2), 213-223.

Slick, D. J., Tan, J. E., Strauss, E. H., & Hultsch, D. F. (2004). Detecting malingering: A survey of experts' practices. *Archives of Clinical Neuropsychology, 19*(4), 465-473.

Strauss, E., Sherman, E. M., Spreen, O., & Spreen, O. (2006). *A compendium of neuropsychological tests: administration, norms, and commentary*. Oxford: Oxford University Press.

Suhr, J. A. & Gunstad, J. (2000). The effects of coaching on the sensitivity and

specificity of malingering measures. *Archives of Clinical Neuropsychology, 15*, 415-424.

Tan, J. E., Slick, D. J., Strauss, E. & Hultsch, D. F. (2002). How'd they do it? Malingering strategies on symptom validity tests. *The Clinical Neuropsychologist, 16*(4), 495-505.

Tombaugh, T. N. (1996). *Test of Memory Malingering (TOMM)*. Toronto, Ont.: Multi-Health Systems.

Wisdom, N. M., Brown, W. L., Chen, D., K., & Collins, R. L. (2012). The use of all three Test of Memory Malingering trials in establishing the level of effort. *Archives of Clinical Neuropsychology, 27*, 208-212.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale–Fourth Edition (WAIS–IV)*. San Antonio, TX: NCS Pearson.

Woods, S. P., Scott, J. C., Conover, E., Marcotte, T. D., Heaton, R. K., Grant, I., & HIV Neurobehavioral Research Center (HNRC) Group. (2005). Test-retest reliability of component process variables within the Hopkins Verbal Learning Test-Revised. *Assessment, 12*(1), 96-100.