The Effects of Third Party Observation on Credible and Non-credible Cognitive Performance: A Simulation Study

A thesis presented to

the faculty of

the College of Arts and Sciences of Ohio University

In partial fulfillment

of the requirements for the degree

Master of Science

Caitlin S. Reese

November 2011

© 2011 Caitlin S. Reese. All Rights Reserved.
This thesis titled
The Effects of Third Party Observation on Credible and Non-credible Cognitive
Performance: A Simulation Study

by

CAITLIN S. REESE

has been approved for
the Department of Psychology
and the College of Arts and Sciences by

Julie A. Suhr
Professor of Psychology

Howard Dewald
Dean, College of Arts and Sciences
Abstract

REESE, CAITLIN S., M.S., November 2011, Psychology

The Effects of Third Party Observation on Credible and Non-credible Cognitive Performance: A Simulation Study

Director of Thesis: Julie A. Suhr

In forensic contexts, judges may court order live observation or video recording of neuropsychological assessments. Professional neuropsychology organizations have issued consensus statements protesting the use of third party observers (TPOs) to preserve the validity of assessments, because TPOs, including video cameras, might inhibit performance on neuropsychological tasks. However, TPO effects are understudied, especially in relation to any possible effect on non-credible responding. The present study investigated TPO effects on the neuropsychological performance of undergraduates with a history of mild TBI. Some participants were asked to simulate impairment. In support of efforts protesting the use of TPO in neuropsychological evaluations, the presence of a video camera impaired visual memory performance among individuals who were instructed to do their best, although a TPO did not differentially affect performance in those asked to simulate impairment. Findings also support the main effect of non-credible effort on memory measures. Taken together, findings further encourage the use of non-credible effort measures in neuropsychological assessment and discourage the use of TPOs in neuropsychological evaluations.

Approved: ________________________________

Julie A. Suhr

Professor of Psychology
Acknowledgments

I would like to acknowledge my advisor, Julie Suhr, Ph.D., and the Ohio University Psychology Department for fostering this opportunity. I would also like to thank my family.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>3</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>4</td>
</tr>
<tr>
<td>List of Tables</td>
<td>7</td>
</tr>
<tr>
<td>Introduction</td>
<td>8</td>
</tr>
<tr>
<td>Method</td>
<td>17</td>
</tr>
<tr>
<td>Participants</td>
<td>17</td>
</tr>
<tr>
<td>Measures</td>
<td>18</td>
</tr>
<tr>
<td>Procedure</td>
<td>19</td>
</tr>
<tr>
<td>Results</td>
<td>22</td>
</tr>
<tr>
<td>Hypothesis one: TPO effects on verbal and visual memory tasks</td>
<td>23</td>
</tr>
<tr>
<td>Hypothesis two: TPO effects on effort measures</td>
<td>24</td>
</tr>
<tr>
<td>Hypothesis three: Effect of non-credible responding</td>
<td>24</td>
</tr>
<tr>
<td>Hypothesis four: Accuracy of detection of non-credible responding</td>
<td>25</td>
</tr>
<tr>
<td>Discussion</td>
<td>27</td>
</tr>
<tr>
<td>References</td>
<td>35</td>
</tr>
<tr>
<td>Appendix A: Instruments</td>
<td>48</td>
</tr>
<tr>
<td>Appendix B: Statistical Analyses</td>
<td>53</td>
</tr>
<tr>
<td>Outlier Analyses</td>
<td>53</td>
</tr>
<tr>
<td>Chi Square Using WMT CNS</td>
<td>54</td>
</tr>
<tr>
<td>Counterbalanced Task Order Analyses</td>
<td>54</td>
</tr>
<tr>
<td>Observation Beliefs</td>
<td>54</td>
</tr>
<tr>
<td>Compliance with Instructions</td>
<td>56</td>
</tr>
<tr>
<td>Control Participants’ Difficulty Ratings</td>
<td>57</td>
</tr>
<tr>
<td>Non-Credible Participants Difficulty Ratings</td>
<td>59</td>
</tr>
<tr>
<td>Appendix C: Other Documentation</td>
<td>61</td>
</tr>
<tr>
<td>Control Instructions</td>
<td>61</td>
</tr>
<tr>
<td>Non-Credible Responding Instructions</td>
<td>62</td>
</tr>
<tr>
<td>Unobserved Consent Form</td>
<td>63</td>
</tr>
<tr>
<td>Form</td>
<td>Page</td>
</tr>
<tr>
<td>---------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Observed Consent Form</td>
<td>65</td>
</tr>
<tr>
<td>Demographics Questionnaire</td>
<td>67</td>
</tr>
<tr>
<td>Debriefing Form</td>
<td>69</td>
</tr>
</tbody>
</table>
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Order of Measure Administration</td>
<td>44</td>
</tr>
<tr>
<td>2</td>
<td>Demographic Characteristics of the Four Groups</td>
<td>45</td>
</tr>
<tr>
<td>3</td>
<td>Descriptive Statistics for Performance on the Verbal Memory D.V.s</td>
<td>46</td>
</tr>
<tr>
<td>4</td>
<td>Descriptive Statistics for Performance on the Visual Memory D.V.s</td>
<td>47</td>
</tr>
<tr>
<td>5</td>
<td>Descriptive Statistics for Controls’ Difficulty Ratings-Delayed Recall</td>
<td>57</td>
</tr>
<tr>
<td>6</td>
<td>Descriptive Statistics for Controls’ Difficulty Ratings-Immediate Recall</td>
<td>58</td>
</tr>
<tr>
<td>7</td>
<td>Descriptive Statistics for Non-Credible Participants’ Ratings-IR</td>
<td>59</td>
</tr>
<tr>
<td>8</td>
<td>Descriptive Statistics for Non-Credible Participants’ Ratings-DR</td>
<td>60</td>
</tr>
</tbody>
</table>
Introduction

Clinical neuropsychologists, often subject to requests for third party observation of litigious neuropsychological evaluations (McSweeny, Becker, Naugle, Snow, Binder, & Thompson, 1998), have long been aware of the third party observer (TPO) effects first examined in the social psychology literature. In general, the pattern of TPO effects has been that simple or well-learned tasks are performed better in the presence of an observer (social facilitation), while difficult or novel tasks are performed worse in the presence of an observer (social inhibition) (McCaffrey, Lynch, & Yantz, 2005). Competing theoretical models have been suggested to explain social facilitation and social inhibition effects, and no mechanism underlying the effects has been agreed upon. The various models to explain third party observation effects have been condensed and organized into drive/arousal theories, social valuation theories, and attention theories (Guerin & Innes, 1984; McCaffrey et al., 2005).

The drive/arousal theories suggest the individual’s tendency to respond is both a function of drive level and the habit of that response for the individual (Zajonc, 1965). Such drive has been described as innate or instinctual (Zajonc, 1965) or the implicit or explicit expectation of evaluation in the presence of an audience (Cottrell, 1968). Increased drive is said to improve performance on easy or well-learned tasks and impair performance on difficult or novel tasks. The social valuation theories, such as the objective self-awareness theory (Wicklund & Duval, 1971) and the self-presentational theory (Bond, 1982) presume individuals actively manage their self-images in the presence of others. When discrepancies between actual and idealized behavior (self-
image) become apparent, individuals are said to perform at a high level to reduce discrepancies, leading to improved performance on easy or well-learned tasks and impaired performance on difficult or novel tasks. Finally, attention theories suggest the task performer experiences attentional conflict in the presence of an observer. This conflict leads the individual to focus on information that is central to the task at hand and lose focus of peripheral details (related to task and observer). Thus, performance on simple or well-learned tasks, which only require focus on what is central to the task, is improved, while performance on difficult or novel tasks, which may require awareness and processing of peripheral cues, is impaired (Baron, 1986). Alternatively, an observer may invoke the controlled processing of information, which could improve performance on tasks that only require automatic processing, but impair performance on tasks that require controlled processing, because the controlled processing resources would be divided between task and observer (Manstead & Semin, 1980).

While no single mechanistic explanation is credited as underlying social facilitation and social inhibition effects, TPO effects began to be documented outside of a strictly experimental setting and within the clinical neuropsychology literature, using tasks commonly administered in a clinical setting. One case study (Binder & Johnson-Greene, 1995) discussed a woman with intractable seizures whose performance on the Portland Digit Recognition Test, a forced-choice measure of recognition memory designed for the purpose of assessing the possibility of malingering, declined in the presence of her mother and improved in her mother’s absence. In another study, using a repeated measures design, undergraduates performed worse on measures of attention,
speed of information processing, and verbal fluency in the presence of a significant other, be it a parent, sibling, close friend, spouse, or partner (Kehrer, Sanchez, Habif, Rosenbaum, & Townes, 2000).

Such findings in the clinical neuropsychology realm raised concerns about the TPO effect leading to diminished validity of evaluations. Other professional concerns about TPOs in a clinical evaluation also began to be expressed. For example, when TPOs witness an evaluation, the security of the measures may be breached. Consequently, critical details of tests could be lost to public access, and testing materials may be misused (Axelrod et al., 2000), such as in the case of coaching test performance (Suhr & Gunstad, 2007). Additionally, the involvement of TPOs violates standardized test administration practices and manuals (e.g. The WMS-III; The Psychological Corporation, 1997). Further, the interpretation of any data collected in these settings, being evaluated against norms drawn from examinations that took place in highly controlled test environments, has the potential to lead to a “spuriously magnified picture of neuropsychological deficit” (NAN, 2000, p. 379). These professional considerations also present implications for ethical practice, considering psychologists are ethically bound to maintain standardized procedures and use test materials with respect to how they have been developed and researched (Ethical Standard 9.02 Use of Assessments), make reasonable efforts to maintain the security and integrity of assessment techniques (Ethical Standard 9.11 Maintaining Test Security), protect against the misrepresentation of their work (Ethical Standard 1.01 Misuse of Psychologists’ Work), and take steps to minimize harm where it is foreseeable and avoidable (Ethical Standard 3.04 Avoiding Harm; Howe
& McCaffrey, 2010). Likely motivated by the aforementioned research findings and professional considerations, the National Academy of Neuropsychology issued a consensus statement (2000) disparaging the use of TPOs during neuropsychological assessments and the American Academy of Clinical Neuropsychology (2001) issued a policy statement discouraging the presence of TPOs in neuropsychological assessments.

In accordance with these findings and position statements, federal courts have mandated TPOs may not be present during neuropsychological (psychological or psychiatric) evaluations. However, state jurisdictions, such as those in Florida and New York, may cite statutes and case law (e.g., cf. Broyles V. Reilly, 1997) permitting the presence of a TPOs during a neuropsychological evaluation. To counteract court orders for third party observation of neuropsychological assessments in such jurisdictions, neuropsychologists have been recommended to provide affidavits describing precedent for excluding TPOs from forensic evaluations (e.g. federal precedent and Ragge v. MCA/Universal Studios), to explain potential issues raised by third party observation, and to provide research findings in support of their position against the use of TPOs.

Following the publication of the aforementioned policy and consensus statements, the body of research addressing neuropsychological evaluations and third party observation expanded noticeably. Such research has evidenced the presence of a neutral graduate student, introduced as an individual assuming the role of a legal representative who wanted to be present for testing, impaired the performance of individuals with a closed head injury on a verbal paired associates delayed recall component of the Wechsler Memory Scale-Revised (Lynch, 2005). Additionally, undergraduate students
who performed in the presence of the evaluator’s supervisor performed worse on the
global and verbal memory subscales of the Memory Assessment Scales than their
unobserved peers (Yantz & McCaffrey, 2005). In another study, undergraduates who
were observed by a female (who sat 1 m behind them, took notes unobtrusively, and did
not make them aware of the reason for her presence) performed worse on a one-category
semantic fluency measure (animals) than their unobserved counterparts (Horwitz &
McCaffrey, 2008). By contrast, the TPO literature has not suggested TPOs facilitate
performance on neuropsychological measures. The absence of such findings is consistent
with the broader TPO literature. Individual exposure to neuropsychological measures
would be expected to be novel and/or more difficult than the simple tasks on which
performance has been facilitated in the TPO literature. As described earlier, performance
on novel or difficult tasks, like neuropsychological measures, is inhibited or impaired in
the presence of TPOs.

To mitigate the professional and ethical considerations of allowing live TPO of
neuropsychological evaluations, McSweeny and colleagues (1998) suggested the use of
audio and video equipment may eliminate TPO effects on individuals. However, some
studies predating their suggestion showed that social facilitation and inhibition effects
remain robust, even when the TPO is indirect or implied. For example, female
undergraduates performed worse on short term recall, and better on long-term recall, of a
paired associates list in the presence of a video camera (Geen, 1973), and the presence of
a closed circuit television or a one-way window improved female undergraduates’
performance on a simple vigilance task (Putz, 1975). In direct response to McSweeney’s
suggestion, this earlier research was extended using clinically relevant measures. Undergraduates tested in an audio tape recording aware group (wherein a tape recorder was placed within close proximity of the participant) performed worse than undergraduates tested in an audio tape unaware group on four Memory Assessment Scales: List Acquisition, Cued Recall, Delayed List Recall, and Delayed Cued Recall (Constantinou, Ashendorf, & McCaffrey, 2002). In another study, when undergraduates completed a battery of neuropsychological tests in the presence of a video camera, observed undergraduates performed worse on Memory Assessment Scales’ List Acquisition, List Acquisition Errors, Number of Learning Trials, Prose Memory, Cued Recall, Delayed List Recall, Delayed Cued Recall, and Delayed Prose Memory (Constantinou, Ashendorf, & McCaffrey, 2005). Thus, research suggests audio and visual recorders should not be used as alternatives to live observation of evaluations because TPO effects remain robust in their presence.

While research investigating TPO effects in neuropsychological assessment has expanded within the past decade, research studies to date have not investigated TPO effects on non-credible effort or malingering. Non-credible effort has been described as “illness behavior or clinical neurological findings that are exaggerated or unrelated to any underlying neurobiological injury or disease” (Albers & Schiffer, 2007, p.312). The use of the descriptor ‘non-credible’ has also been adopted because “it does not commit an expert witness to any specific psychological motivation for the production of the behavior or findings” (Albers & Schiffer, 2007, p.312). That is, multiple conditions could be culpable for producing such exaggerated behaviors, including conversion disorder,
factitious disorder, psychological pain disorder, somatoform disorder, and blatant attempts to mislead clinicians (i.e. malingering). Forensic clinicians may likely find it impossible to distinguish among these syndromes and motivations behind non-credible effort (Albers & Schiffer, 2007), but non-credible effort in general should be considered a reality for which to assess during neuropsychological assessment, especially in the forensic arena.

Larrabee (2007) suggested the base rate of non-credible effort in forensic settings may exceed 50% (based on a survey of American Board of Clinical Neuropsychology diplomats who reported on their forensic work), and determined the base rate of malingering among personal injury litigants alleging mild head injury was 38.5% (41.24% when adjusted for referral source). Larrabee (2003) reviewed 11 studies, encompassing 1,363 consecutively evaluated mild traumatic brain injury litigants and found 40% of them performed non-credibly based on formal assessment. Miller, Boyd, Cohn, Wilson, and McFarland (2006) found that 54% of Social Security disability applicants failed formal measures of non-credible effort. Mild TBI cases also represent the majority of patients who consult with forensic neuropsychological practitioners (Bender, 2008).

While studies have examined TPO effects in individuals who presumably are motivated to perform their best, it is unknown how TPO affects the performance of individuals who intend on feigning cognitive impairment. Preliminary research, presented in the form of a poster (Condit, Mittenberg, Colden, & Shapiro, 2008) found that the three non-credible effort indices of the Word Memory Test (WMT) accurately classified 100% of video-recorded feigning participants, and the WMT consistency index classified
statistically significantly more video-recorded feigners than non-video recorded feigners, suggesting that the presence of the video camera enabled the better classification of non-credible effort. However, they found no difference in performance between observed and unobserved feigning on the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III) subtests. It is important to note that many non-credible effort measures, including the WMT, are extremely easy cognitive tasks (although they may appear difficult), relative to other cognitive tasks that accurately assess neuropsychological performance. However, when one’s objective is to realistically appear impaired on a task, it is difficult to know whether non-credible effort tasks are in fact “easy.” With knowledge of the aforementioned social psychological theories, and their explanations for how TPOs inhibit performance on novel, difficult tasks, more research is necessary to better understand what it means to willfully inhibit effort during third party observation.

The present study was conducted to examine the effect of the presence of a video camera on neuropsychological performance among individuals who intend to perform with their best effort, as well as those who intend on feigning cognitive impairment. It was hypothesized that observed controls would perform worse than unobserved controls on memory tasks due to TPO inhibition effects on these more difficult tasks, although it was not expected that TPO would affect performance on measures of non-credible effort in this group, given the simplicity of measures of non-credible effort. Further, it was expected that, collapsed across observation, those feigning cognitive impairment would perform worse than controls on both memory and non-credible effort measures. Exploratory analyses were also completed to investigate possible effects of the indirect
third party observation on the detection of non-credible effort in those purposely feigning cognitive impairment.
Method

Participants

Following approval from the human subjects Institutional Review Board, 112 undergraduate students were recruited from introductory psychology classes after they completed an online screening measure. Participants who endorsed being between the ages of 18 and 25 years old, reported having a history of mild TBI/concussion lasting between a few seconds and 20 minutes, and denied engaging in substance abuse or having a history of psychological, neurological, or learning disorder were invited to participate in the study. Ten students were excluded from the analyses because they endorsed having exclusionary diagnoses on the study’s demographics questionnaire at the time of their participation. Three of these students endorsed having an anxiety disorder, one student endorsed having anxiety and depression, two students endorsed having “ADD,” two students endorsed having epilepsy, and two students indicated they had a diagnosis, but did not specify which diagnosis. Additionally, one participant was excluded because the student accidentally completed the post-task form, instead of the task rating form, after the first task, and it is possible the participant prematurely saw the question asking, do you believe you were videotaped? Another participant was excluded because the participant explained he did not follow the instructions to feign impairment because he “would never do that.” The remaining participants identified as follows: 45 women and 55 men, 91 Caucasian/non-Hispanic, 3 African American, 2 Asian/Pacific
Islander, and 4 two or more races. Participants’ class standing ranged from freshman to senior (approximately 71% of the participants were freshmen).

Participants were randomly assigned to one of four testing conditions: Observed non-credible effort (n=26); unobserved non-credible effort (n=22); observed control (n=24); and unobserved control (n=28) (see Table 2 for demographic information).

Measures

The paramount instruments used in the present study included four tests: Word Memory Test, Rey Auditory Verbal Learning Test, Continuous Visual Memory Test, and Rey Complex Figure Test, encompassing 9 dependent measures. These tests, including relevant dependent measures for the present study are briefly described below (additional psychometric information can be found in Appendix A).

The Word Memory Test (WMT; Green & Astner, 1995) is a well-validated computerized task commonly used to measure non-credible responding in neuropsychological testing. The present study utilized the immediate recognition index score as a non-credible effort measure. The WMT User’s Manual stipulates a “clear fail” is evident when someone scores at or below 82.5% correct (Green & Astner, 1995, p. 9).

The Rey Auditory Verbal Learning Test (RAVLT; Rey, 1964) is a well-validated measure of list learning, immediate memory, delayed recall and recognition, and susceptibility to interference. Dependent variables used in the present study included the immediate memory trial, 30-minute delayed recall trial, 30-minute delayed recognition trial, and a learning over trials index.
The Continuous Visual Memory Test (CVMT; Trahan & Larrabee, 1997) is a well-validated visual recognition memory test. Dependent variables used in the present study were the CVMT acquisition total, delayed recognition task, and d-prime score.

The Rey-Osterrieth Complex Figure Test (Osterrieth, 1944) requires participants to copy a complex figure. After a delay, including intermittent verbal tasks, they are asked to recopy the figure from memory. The test is well validated as a measure of visuospatial constructional ability and visual, perceptual, motor, and episodic memory. The CFT total recall score was included as a dependent variable in the present study.

**Procedure**

In order to enhance the ecological validity of the study, each participant was administered the aforementioned measures embedded within a full neuropsychological battery of 12 clinical instruments (brief explanations of the additional instruments are included in Appendix A). The order of measure administration was consistent with one of two predetermined, counterbalanced orders (detailed in Table 1) and randomly assigned to each participant.

Two experimenters were required to test each participant. One experimenter uncovered the randomized test condition: observed non-credible; unobserved non-credible; observed control; unobserved, control (note: any reference to non-credible responder refers to simulated non-credible effort, see below). The order of measure administration (one or two) to be run for that slot’s participant was also uncovered at that time. That experimenter set up the testing room in accordance with the condition and order of the measures. If the condition was an observed condition, a video camera was set
up in plain sight, set up off to the side, oriented at a right angle to the plane of interaction between the experimenter and examinee.

In a separate room, the first experimenter reviewed the consent process with the participant. Participants were informed they were providing consent to participate in voluntary research exploring performance on a variety of cognitive tests in different testing environments. Observed participants’ consent forms differed from unobserved participants’, as they were notified they may or may not be videotaped (see Appendix C for consent forms). The first experimenter also provided instructions to all participants. Best effort condition participants were asked to do their best, while non-credible condition participants were asked to realistically portray (i.e. simulate) what they thought TBI symptoms might “look like” during neuropsychological testing. The latter were not instructed how to feign impairment; they were supplied a lawyer’s website supplement (www.braininjury.com/symptoms.shtml) describing symptoms of brain injury and told there was no wrong answer (we reasoned individuals intending to feign cognitive impairment on a neuropsychological evaluation might access the Internet for ideas/information). The non-credible participant instructions also informed participants it would be important not to act overtly obviously to keep from being detected by the second experimenter (full directions are also included in Appendix C) and rate (on a 5-point scale) how difficult feigning would be.

Participants were then introduced to the second experimenter (who was unaware of their assigned condition and instructions). The second experimenter administered the neuropsychological battery; he or she was either the primary investigator or a research
assistant trained by a primary investigator or under the primary investigator’s supervision (researchers alternated between first and second experimenter depending upon schedules and availability). The second experimenter asked participants in observed conditions to complete their demographics questionnaire (located in Appendix C) while the second experimenter stated he or she had to finish setting up the video camera. Participants were not aware of the fact the experimenter did not press record on the video camera (participants were not actually recorded). During the battery, non-credible participants completed a rating form to rate how difficult each task was and how difficult it was to feign cognitive impairment on that task. Control participants completed a rating form to rate how difficult each task was and how long they believe that task took to complete. Following the administration, participants completed a post-task form with a compliance check, asking them to rate how successful they were in following their instructions, and a manipulation check asking them if they believed they were being videotaped. They were then debriefed and assigned experimental credit. Most participants were also paid $10; 81 participants were paid for their participation and 32 participants were not paid. Participants were paid after the first ten weeks of recruitment failed to recruit many participants. Additionally, financial resources were no longer available to pay participants during the last ten weeks (approximately) of recruitment.
Results

Participant age, education, ethnicity, handedness, and sex were evenly distributed among groups (see Table 2 for demographic information). Checks for the normality of the data revealed that non-credible responders’ performance generally fit a normal distribution on all measures (one non-credible responder was an outlier on the AVLT learning over trials scale, but we asked non-credible responders to feign impairment, so it was not surprising someone feigned to an extreme extent). Control participants’ performance fit a normal distribution on many measures, but the AVLT 30-minute recall scale was negatively skewed, the AVLT 30-minute recognition scale was negatively skewed and kurtotic, and the CFT total recall scale was negatively skewed. In addition, potential outlier scores were identified as follows: 2 controls performed as outliers on the AVLT 30-minute recall scale, 3 controls performed as outliers on the AVLT 30-minute recognition scale, 3 controls performed as outliers on the CFT total recall scale, and 4 controls performed as outliers on the CVMT acquisition total scale. Outliers were always in the direction of impairment on the dependent measure. No control participants were identified as outliers on more than one measure. Of the two control participants who failed the effort measure (WMT IR), one was in the observed group and one was in the unobserved group; both of these individuals performed as outliers only on the AVLT 30-minute recognition scale. Removal of outliers normalized the data for all but the AVLT 30-minute recognition distribution, which remained negatively skewed.

In order to address these findings, data were analyzed with and without outlier scores in the models, and results from both analyses will be compared and described. We
opted not to use nonparametric statistics when the outliers were retained, as many data analysts believe MANOVAs are extremely robust against violations of normality (Ito, 1980; vonEye & Bogat, 2004) and result in fairly similar type I error rate and power (Finch, 2005).

**Hypothesis one: TPO effects on verbal and visual memory tasks**

It was hypothesized observed control participants would perform worse than unobserved control participants on the verbal and visual memory tasks, and separate Multivariate Analyses of Variance models (MANOVA) were conducted to test these hypotheses. Observed and unobserved controls did not perform differently on verbal memory tasks, $A = .958, F(4, 46) = .502, p = .734$. Effect sizes were small (RAVLT IR $d = .10$, RAVLT 30-minute recall $d = .18$, LOT $d = .31$, and RAVLT 30-minute recognition $d = .11$). Results were similar with outliers removed, $A = .96, F(4, 34) = .364, p = .832$.

However, observed controls performed significantly worse than unobserved controls on visual memory tasks, $A = .73, F(4, 46) = 4.19, p = .006$. Using Bonferroni correction (.05/4), follow-up ANOVAs revealed observed controls performed worse than unobserved controls on the CVMT d-prime task, $F(1,49) = 4.63, p = .000, d = 1.08$, and the CVMT acquisition total, $F(1,49) = 413.0, p = .001, d = .95$. The groups did not perform differently on CFT recall, $F(1,49) = 6.73, p = .663, d = .12$; or CVMT delayed recognition, $F(1,49) = 2.22, p = .252, d = .33$. Results, with the same pattern of follow-up tests, remained the same even after the removal of statistical outliers, $A = .72, F(4, 34) = 3.36, p = .020$. 
Hypothesis two: TPO effects on effort measures.

In order to examine whether the two control groups would differ in failure on a non-credible effort measure, their performance was compared using chi square analysis. There was no difference between observed and unobserved control participants on WMT IR failure, $\chi^2(1) = .01, p = .91$; 4.2% of observed controls and 3.6% of unobserved controls failed the WMT IR.

Hypothesis three: Effect of non-credible responding on verbal and visual memory tests.

Additional hypotheses posited, collapsed across observation, non-credible responders would perform worse on verbal and visual memory tests than controls. Using Pillai’s trace (Box’s test was significant), the verbal memory MANOVA indicated there was a significant effect of non-credible effort on verbal memory measure performance, $V = .37, F(4, 94) = 14.05, p = .000$. Follow-up ANOVAs indicated non-credible responders performed worse on RAVLT immediate recognition, $F(1,97)=42.95, p=.000, d=1.31$; RAVLT 30-minute recall, $F(1,97)=48.48, p=.000, d=1.39$; RAVLT Learning Over Trials, $F(1,97)=16.37, p=.000, d=.81$; and RAVLT 30-minute recognition, $F(1,97)=52.39, p=.000, d=1.44$. The significant effect of non-credible effort on verbal memory remained after removing the outliers, $V = .42, F (4, 82) = 14.87, p < .001$, with the same pattern of results seen in follow-up tests.

The visual memory MANOVA also indicated non-credible participants performed worse than control participants on visual memory measures, $V = .38, F (4, 93) = 14.10, p = .000$. Follow-up ANOVAs indicated non-credible participants performed worse on the
CFT recall, $F(1,96)=29.89, p=.000, d=1.10$; the CVMT d-prime task, $F(1,96)=28.60, p=.000, d=1.07$; the CVMT acquisition total, $F(1,96)=36.83, p=.000, d=1.21$; and the CVMT delayed recognition, $F(1,96)=45.37, p=.000, d=1.35$. The effect remained with the removal of outliers, $V = .43, \chi^2 (4, 81) = 15.46, p < .001$, with the same pattern of results seen in follow-up tests.

**Hypothesis four: Accuracy of detection of non-credible responding**

We also expected non-credible responders would be detected as such, based on WMT IR failure, more often than control participants. Consistent with that hypothesis, the proportion of WMT IR index failure was not evenly distributed between groups, $\chi^2 (1) = 57.97, p < .001$. In all, 37 of 47 (78.7%) non-credible responders failed the WMT IR and 2 of 52 (3.8%) control participants failed the WMT IR. An ROC curve analysis showed WMT IR was able to effectively discriminate non-credible performers from controls, AUC=.93. This finding is consistent with existing literature supporting the use of the WMT as a psychometrically sound credibility measure.

**Hypothesis Five: Effects of TPO on Non-Credible Performance**

Two exploratory MANOVAs were conducted to explore the effects of the presence of an indirect TPO on performance within the group asked to simulate impairment. Using Wilks' lambda, observed non-credible participants and unobserved non-credible participants did not perform differently on the verbal memory measures, $\Lambda = .93, F (4, 43) = .77, p = .552$. The effect sizes were generally small (RAVLT IR $d=.34$, RAVLT 30-minute recall $d = .08$, LOT $d=.36$, and RAVLT 30-minute recognition $d =.25$). Similarly, there was no difference on the visual memory measures, $V = .08, F (4,$
42) = .95, \( p = .446 \), and the effect sizes were small (CVMT total \( d = .16 \), CVMT d-prime \( d = .04 \), CVMT delayed recognition \( d = .22 \), and CFT total \( d = .30 \)). Furthermore, there was no difference between observed and unobserved non-credible performance on WMT IR failure, \( \chi^2(1) = .15, \ p = .70 \); 80.8% of observed non-credible responders and 76.2% of unobserved non-credible responders failed the WMT IR.
Discussion

The present study sought to extend prior research showing third party observation impairs credible neuropsychological performance, even when the TPO is an indirect observer such as a video camera. In the present study, participants with a history of mild TBI, who were instructed to do their best, performed worse on measures of visual memory than their counterparts in the observed condition. This finding is of value because it provides additional evidence that using a video camera to observe neuropsychological evaluations negatively impacts the performance of individuals who intend on performing credibly during the assessment. Furthermore, these findings expand upon those of previous studies, because previous research has not revealed a TPO effect on visual memory (e.g. Yantz & McCaffrey, 2005). In addition, the effect was not merely due to a few participants exhibiting abnormally low performance and driving down the mean, as the difference remained even after outlier scores were removed.

While it was hypothesized observed control participants would also perform worse on verbal memory measures, contrary to the existing literature, the results were not statistically significant, and the effect sizes were small. In fact, the sample size would have to increase to at least 170 participants per cell (on learning over trials) or more (immediate recall = 1,800 per cell; 30-minute recognition = 1,560 per cell; and 30-minute recall = in excess of 9,500 per cell) to potentially detect a difference on verbal memory measures. The fact that no TPO effect on verbal memory measures was apparent in the present study is counterintuitive, considering previous research has found both direct and indirect TPO effects on various verbal memory measures, including delayed recall of
verbal paired associates (Lynch, 2005) and various verbal memory subscales of the Memory Assessment Scales (Constantinou et al., 2002; Constantinou et al., 2005; Yantz & McCaffrey, 2005), sometimes with large effect sizes. Like the RAVLT, the Memory Assessment Scales contain a list learning task, a list recall task, and a delayed recall task. The RAVLT presents 15 unrelated words 5 times, then presents an interference list, and presents recall and recognition trials following a 30 minute delay. By comparison, the MAS presents 12 common words, each belonging to 4 categories, and the list is presented only until all 12 words are successfully recalled on a test trial (up to six times maximum). Following delays, examinees are asked to recall the words and categories with and without cues. Constantinou and colleagues (2002; 2005) found the unobserved (via audio recorder or video camera) groups scored higher on initial MAS list acquisition. It is possible significant differences in list acquisition translated into differences in encoding (i.e. someone who recalled 12 words on the second trial had less opportunity to hear and encode the words subsequently than someone who heard the word list all six times) and later in recall. These potential differences would less likely be seen on the RAVLT because everyone hears the word list the same number of times (five), and has the same opportunity for encoding. This difference could make the task seem potentially easier than the MAS, accordingly decreasing its vulnerability to social inhibition effects. These results also suggest the RAVLT may be more robust against TPO effects. It is not likely that the absence of statistical differences is related to sub-optimal effort (e.g. on the part of the unobserved control group), because one participant from both the observed and unobserved control groups failed the effort measure (and might similarly impact their
respective overall means). Also, results did not change when their outlier scores were
removed from analyses.

It is also possible that our findings were only apparent on the visual memory tasks
because they were judged to be more difficult than the verbal memory tasks.
Supplementary t-tests indicated, on average, observed controls rated the CVMT
acquisition total as more difficult than unobserved controls (and prior research shows the
presence of a TPO impairs performance on difficult tasks); whereas, they rated the WMT
IR and the AVLT IR as equally difficult (see Appendix B). This specific between group
difference for the CVMT acquisition total could account for a TPO effect specific to the
CVMT acquisition total and d-prime scores (as the d-prime score is derived from
components of the CVMT acquisition total).

Given the robust effect sizes of non-credible effort on neuropsychological test
performance in the existing literature, it is not surprising we found that individuals who
intend on feigning impairment perform worse on verbal and visual memory tasks than
individuals instructed to perform with their best effort. Non-credible responders were not
instructed how to feign cognitive impairment; rather, they were supplied with a head
injury supplement downloaded from a disability lawyer's website, and they made
individual decisions as to how they might feign cognitive impairment. We consider the
use of the website supplement to be a realistic proxy of an initial step (i.e. searching the
Internet) someone seeking an external incentive in a forensic setting might take. In
addition, the use of individuals who reported a previous TBI history is an additional
strength to the simulator design. Participants who were instructed to feign cognitive
impairment (in a manner to avoid getting caught by the examiner, who was unaware of group status) indeed performed worse on memory measures than those instructed to perform with their best effort. However, they were also detected as non-credible on the WMT in 78.7% of cases. In a clinical evaluation situation, individuals who research how to feign TBI sequelae (in a manner to avoid getting caught by a clinician) would likely also perform significantly poorly. Unfortunately, without the use of a non-credible effort measure (such as the Word Memory Test), feigned performance in a clinical setting might unknowingly and invalidly be interpreted as impaired performance and ability, reflective of brain damage, and worthy of external compensation.

It is of note that, although the WMT IR scale was able to detect non-credible responders overall, it did not disproportionately categorize observed and unobserved non-credible responders. Using a different effort measure, the WMT CNS scale, Condit and colleagues’ (2008) found that the WMT CNS accurately classified 100% of the non-credible responders observed by video camera, while it classified significantly fewer (80%) unobserved non-credible responders. Condit and colleagues’ findings essentially suggest the presence of a TPO could allow for the detection of more non-credible responders. We reran the chi square analyses using the WMT CNS, and there was no difference between observed and unobserved non-credible performance on WMT CNS (Appendix B). The fact these findings were not replicated here hopefully discourages any suggestion that the use of a TPO during neuropsychological assessment promotes the detection of non-credible responding, or any consequent argument to include a video camera in the assessment environment.
Following exploratory analyses, we did not observe a TPO effect on observed non-credible responders; that is, observed and unobserved non-credible responders did not perform differently on verbal and visual memory measures. In addition, the small effect sizes suggest the absence of differences between observed and unobserved non-credible responders was not merely due to sample size. In this case, the sample size would have to increase to between 130 and 1,550 participants per cell on the verbal memory subscales (Immediate Recall=150 per cell; 30-Minute Recall = 1,550 per cell; Learning Over Trials=130 per cell; and 30-Minute Recognition = 290 per cell) and between 171 and in excess of 9,500 on the visual memory subtests (CVMT acquisition total=660 per cell; CVMT d-prime= > 9,500; delayed recognition =300 per cell; and CFT total = 171 per cell) to detect differences between the observed and unobserved non-credible responders. As shown in appendix B, observed and unobserved non-credible responders rated the difficulty of the tasks equally. There was also no difference in their perceived success in following their instructions to feign impairment. Perhaps we would have found evidence of a TPO effect if the observed non-credible responders had perceived a task as being more difficult, or if they had felt less successful feigning cognitive impairment than their unobserved counterparts (Appendix B).

This study presents a number of potential limitations. For example, one potential limitation of the study is that it did not include an observer manipulation check until the study concluded. Initially, a break was given half way through the administration, and the intention of the break was to remind participants of the observation (e.g. “You can take a break while I check the tape in the camera”). After two students commented (following
their participation) that they did not believe they were videotaped because “the tape is a 6-hour videotape, why would it need to be checked after an hour?” this manipulation check was removed from the script. An additional manipulation check, “Do you believe you were videotaped?” was retained and included on a form given at the end of the study. While some participants reported they did not believe they were videotaped, it is also possible they reported, “no” to appear sophisticated, intellectual, or in an effort to fulfill some demand characteristic (e.g. negative participant role). However, after removing individuals who either did not believe they were videotaped when in the presence of a video camera, or believed they were videotaped in the absence of the video camera, the reported results did not change (Appendix B).

Regarding additional potential limitations, each participant completed the study in between 2 and 2.5 hours, depending on how long it took them to complete the tasks (many of the non-credible responders took extended periods of time to do so). It is possible the lengthy administration time might render some of the tasks susceptible to fatigue related differences. To account for any possible fatigue effects, we counterbalanced the order of task administration. There was no main effect of task administration order collapsed across group (Appendix B).

Some may also question the generalizability of this study’s results because the non-credible responders were simulating. Rogers (1988) critiqued simulation studies because, while they boast a high degree of experimental control and internal validity, their external, ecological validity is limited. Inman and Berry (2002) improved upon simulator designs by including detailed malingering instructions, manipulation checks, an
incentive, a control group, multiple indices, appropriate analyses, a test administrator who is blind to the experimental condition, malingerers and controls with a history of head injury, clinically established cutting scores, and a battery of tests that mimics the experience of a standard neuropsychological assessment. For the most part, the same simulator design was used during this study, with the exception of using an incentive. A number of studies (Bernard, 1990, Erdal, 2004, and Weber, 2008) have found that simulating malingerers (typically college undergraduates) offered a monetary incentive do not evade malingering detection at a higher rate than simulators who do not receive monetary incentives, and they sometimes simulate non-credible responding more flagrantly than undergraduate simulators who were not offered an incentive (Erdal, 2004). Also defending the use of a design using simulators, it would be unethical to conduct this nature of research on actual patients, given existing evidence TPOs may affect their performance.

In summary, the present results provide mixed evidence that simulating non-credible responders perform worse than individuals intending to do their best on memory measures. It further suggests allowing TPOs in clinical evaluations may lead to invalid conclusions pertaining to well-intending patients’ visual memory skills. These results should be considered useful to further bolster affidavits against the use of TPOs in jurisdictions wherein third party observation of neuropsychological evaluations continues to be court ordered.

Future research may consider whether or not various TPOs (direct, indirect, known, unknown, supportive, aversive, etc.) impose a universal effect or varying degrees
of impairment on neuropsychological measures. Such research might be conducted through the use of measures with alternate forms and/or parallel tasks. In addition, while research in the realm of TPOs and neuropsychology has not focused on elucidating the theoretical process that induces TPO effects, future research may utilize measures of evaluation apprehension in an effort to begin to do so. Perhaps such measures could be administered immediately after a participant was made aware of a TPO and subsequently during the battery; differences in degree of evaluation apprehension, between observed and unobserved individuals, could ultimately support arguments favoring drive/arousal or social valuation theories to explain TPO effects. Considering many models of supervision include a TPO (supervisor), exploring whether a particular type of TPO minimizes TPO effects, or whether including a measure of evaluation apprehension could cue when a patient’s results should be interpreted with caution (based on degree of evaluation apprehension), could be of training and clinical value.
References


Table 1

*Order of Measure Administration*

<table>
<thead>
<tr>
<th>Order One</th>
<th>Order Two</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMT (Immediate Recognition)</td>
<td>WMT (Immediate Recognition)</td>
</tr>
<tr>
<td>CFT</td>
<td>TMT A&amp;B</td>
</tr>
<tr>
<td>COWAT</td>
<td>Grooved Pegboard</td>
</tr>
<tr>
<td>CVMT (Acquisition)</td>
<td>Grip Strength Test</td>
</tr>
<tr>
<td>WTAR</td>
<td>Digit Span Subtest</td>
</tr>
<tr>
<td>WMT (Delayed Recognition)</td>
<td>Finger Tapping Test</td>
</tr>
<tr>
<td>CFT (Delayed Recall)</td>
<td>WMT (Delayed Recognition)</td>
</tr>
<tr>
<td>RAVLT (Learning/Immediate Recall)</td>
<td>CVMT (Acquisition)</td>
</tr>
<tr>
<td>CVMT (Delayed Recognition)</td>
<td>WTAR</td>
</tr>
<tr>
<td>CVMT (Visual Discrimination Task)</td>
<td>RAVLT (Learning/Immediate Recall)</td>
</tr>
<tr>
<td>TMT A&amp;B</td>
<td>CFT</td>
</tr>
<tr>
<td>Grooved Pegboard</td>
<td>COWAT</td>
</tr>
<tr>
<td>Grip Strength Test</td>
<td>CVMT (Delayed Recognition)</td>
</tr>
<tr>
<td>Digit Span Subtest</td>
<td>CVMT (Visual Discrimination Task)</td>
</tr>
<tr>
<td>Finger Tapping Test</td>
<td>CFT (Delayed Recall)</td>
</tr>
<tr>
<td>RAVLT (30-Minute Recall/Recognition)</td>
<td>RAVLT (30-Minute Recall/Recognition)</td>
</tr>
<tr>
<td>MMPI-2-RF</td>
<td>MMPI-2-RF</td>
</tr>
<tr>
<td>RAVLT (60-Minute Recall/Recognition)</td>
<td>RAVLT (60-Minute Recall/Recognition)</td>
</tr>
</tbody>
</table>

*Note.* The effort measure, visual memory measures, and verbal memory measure included in analyses are noted in boldface. The test acronyms are as follows: Complex Figure Test (CFT), Controlled Oral Word Association Test (COWAT), Continuous Visual Memory Test (CVMT), Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF), Rey Auditory Verbal Learning Test (RAVLT), Trail Making Test (TMT), Wechsler Test of Adult Reading (WTAR), and Word Memory Test (WMT).
Table 2

Demographic Characteristics of the Four Groups

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Age</th>
<th>Education</th>
<th># of Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed Poor Effort</td>
<td>26</td>
<td>19.31(1.12)</td>
<td>1.69(1.01)</td>
<td>18</td>
</tr>
<tr>
<td>Unobserved Poor Effort</td>
<td>22</td>
<td>18.95(1.13)</td>
<td>1.27(.70)</td>
<td>13</td>
</tr>
<tr>
<td>Observed Best Effort</td>
<td>24</td>
<td>18.83(.82)</td>
<td>1.29(.55)</td>
<td>10</td>
</tr>
<tr>
<td>Unobserved Best Effort</td>
<td>28</td>
<td>19.00(1.15)</td>
<td>1.36(.56)</td>
<td>14</td>
</tr>
</tbody>
</table>

Note. Age and education are included in mean (S.D.) format. Education was coded as follows: first year (1), second year (2), third year (3), fourth year (4), fifth or above (5).
Table 3

Descriptive Statistics for Performance on the Verbal Memory Dependent Variables

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>IR</th>
<th>30-Recall</th>
<th>30-Recog</th>
<th>LOT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs. Con.</td>
<td>23</td>
<td>12.04 (2.23)</td>
<td>11.30 (2.53)</td>
<td>14.13 (1.22)</td>
<td>19.43 (6.93)</td>
</tr>
<tr>
<td>Unobs. Con.</td>
<td>28</td>
<td>11.82 (2.14)</td>
<td>11.75 (2.44)</td>
<td>14.25 (1.04)</td>
<td>21.54 (6.79)</td>
</tr>
<tr>
<td>Obs. Non-Cred</td>
<td>26</td>
<td>8.42 (3.78)</td>
<td>7.46 (3.73)</td>
<td>11.35 (2.91)</td>
<td>15.92 (6.75)</td>
</tr>
<tr>
<td>Unobs. Non-Cred</td>
<td>22</td>
<td>7.09 (3.98)</td>
<td>7.14 (4.20)</td>
<td>10.59 (3.14)</td>
<td>13.18 (8.59)</td>
</tr>
<tr>
<td>Non-Cred</td>
<td>48</td>
<td>7.81 (3.88)</td>
<td>7.31 (3.91)</td>
<td>11.00 (3.01)</td>
<td>14.67 (7.69)</td>
</tr>
<tr>
<td>Control</td>
<td>51</td>
<td>11.92 (2.16)</td>
<td>11.55 (2.47)</td>
<td>14.20 (1.11)</td>
<td>20.59 (6.87)</td>
</tr>
</tbody>
</table>

Note. Scores are included in mean (S.D.) format. The RAVLT is the Rey Auditory Verbal Learning Test.
Table 4

Descriptive Statistics for Performance on the Visual Memory Dependent Variables

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>CVMT Total</th>
<th>CVMT d-prime</th>
<th>CVMT DRT</th>
<th>CFT Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed Con.</td>
<td>23</td>
<td>74.17 (6.60)</td>
<td>1.66 (.59)</td>
<td>4.65 (1.19)</td>
<td>23.91 (4.77)</td>
</tr>
<tr>
<td>Unobs. Con.</td>
<td>28</td>
<td>78.82 (8.39)</td>
<td>2.26 (.54)</td>
<td>5.07 (1.36)</td>
<td>24.64 (6.71)</td>
</tr>
<tr>
<td>Observed N-C</td>
<td>26</td>
<td>66.85 (10.46)</td>
<td>1.19 (.66)</td>
<td>2.96 (1.73)</td>
<td>18.60 (5.86)</td>
</tr>
<tr>
<td>Unobs. N-C</td>
<td>21</td>
<td>64.95 (12.67)</td>
<td>1.16 (1.09)</td>
<td>2.57 (1.83)</td>
<td>16.81 (5.98)</td>
</tr>
<tr>
<td>Non-Credible</td>
<td>47</td>
<td>66.00 (11.41)</td>
<td>1.18 (.86)</td>
<td>2.79 (1.77)</td>
<td>17.80 (5.92)</td>
</tr>
<tr>
<td>Control</td>
<td>51</td>
<td>76.73 (7.91)</td>
<td>1.99 (.63)</td>
<td>4.88 (1.29)</td>
<td>24.31 (5.87)</td>
</tr>
</tbody>
</table>

Note. Scores are included in mean (S.D.) format. The test acronyms are as follows: Complex Figure Test (CFT) and Continuous Visual Memory Test (CVMT). Control is shortened to “Con.” and Unobserved is shortened to Unobs. Non-credible is shortened to “N-C.”
Appendix A: Instruments

Word Memory Test- Failure on the WMT effort measures is associated with a general over-reporting of symptoms (Gervais et al., 2001). The WMT has been found to be psychometrically sound. Green (2003) reported high internal consistency between IR and DR ($r=.88$), Effort measures (IR and DR) correlated highly with one another on initial testing ($r=.87$) and retesting after one year or more ($r=.94$). This discussion of WMT psychometrics is limited because the immediate recognition index was of interest to this study.

Rey Auditory Verbal Learning Test- Over one-year intervals, test-retest reliability has been shown to be marginal to adequate. The fifth trial of the delayed-recall trials was among the more reliable scores ($r$ values about .60 to .70; Uchiyama et al., 1995). Regarding concurrent validity, the AVLT loads with other verbal memory measures and general verbal ability measures. It does not load with motor or visuospatial measures, and it strongly correlates with other verbal memory measures (Schmidt, 1996, p.48). It has been shown to discriminate between normal and mixed neurological groups (Powell et al., 1991), and different patient groups perform differently on the test (Schmidt, 1996, p. 69).
Continuous Visual Memory Test- The three scores that are most useful for analyzing visual memory include acquisition total score, the d-prime score (which represents the z-score for false alarms subtracted from the z-score for hits), and the delayed recognition score. In normal subjects, and those with a history of severe closed head injury, the CVMT has a high degree of inter-item reliability for both recurring ($r=.80$, $r=.98$) and nonrecurring test items ($r=.98$, $r=.90$), respectively. Additionally, test-retest reliability, following a seven-day interval ranged between .76 and .85 (Winer, 1971). Regarding construct validity, the d-prime score has loaded on a general cognitive factor, and a clear visual memory factor has been defined by the CVMT Delayed Recognition Task (.68; Trahan & Larrabee, 1984). Trahan & Larrabee (1984) concluded that the d-prime score is more closely associated with cognitive processing and encoding factors, while the CVMT Delayed Recognition Task appears to define an independent dimension of visual memory functioning.

Complex Figure Test- Fasteneau and colleagues (1996) treated each detail of the CFT as an item and subjected scores to split half and coefficient alpha procedures. They found that both split-half and coefficient alpha reliabilities were above .60 for the copy condition and above .80 for the recall conditions. Regarding test-retest reliability, in a sample of healthy adults, after a six-month interval, test-retest reliability was found to be $r=.76$ for immediate recall, $r=.89$ for delayed recall, and $r=.87$ for recognition total (Meyers & Meyers, 1995). Data from correlational and factor analytic studies bolster the
construct validity for the CFT as a measure of visuoconstructional ability (copy) and memory (recall and recognition; Meyers & Meyers, 1995; Shorr et al., 1992).

The measures described below are not germane to this study’s hypotheses; however they were included in the full neuropsychological battery to enhance the ecological validity of the study.

The Wechsler Test of Adult Reading- (WTAR; Wechsler, 2001) is a word-reading test that involves pronouncing 50 irregularly spelled words. The WTAR is scored in terms of the number of correctly pronounced words.

The Trail Making Test, Parts A and B- (TMT; Reitan & Wolfson, 1995) is a timed test which measures attention, mental flexibility, sequencing ability, visual scanning, and motor function. In part A of the TMT, participants are instructed to draw lines to numerically sequence 25 numbers. In part B of the TMT, participants are instructed to draw lines sequencing 25 numbers and letters in numerical and alphabetical order, respectively. Participants’ performance is based on completion time.

The Controlled Oral Word Association Test- (COWAT; Benton, Hamsher, & Sivan, 1994) assesses the neuropsychological construct of verbal association fluency. In the COWAT, participants are instructed to name as many words that begin with a certain letter (e.g. A, F, or S) as quickly as possible. Participants are allowed one minute to complete each trial of the test, and performance is based on word totals.
The Grooved Pegboard Test- (Lafayette Instruments, Lafayette, IN) is administered to evaluate the dexterity of the participant’s hands. The test consists of a board with 25 small holes in which 25 small pegs are to be placed as quickly as possible. Performance is based on completion time.

The Finger Tapping Test- (Shimoyama, Ninchoji, & Uemura, 1990) measures the impact of movement disorders (e.g. dyskinesia and tremor) on extremity function. The measure also assesses for subtle motor impairment, cognitive impairment, and laterality of brain lesions. Participants, one hand at a time, alternately tap two buttons (that are separated by 30 cm) over the course of one minute. The buttons are attached to counters that yield the participant’s score.

The Wechsler Adult Intelligence Scale – IV Digit Span Subtest- (Wechsler, 2008) requires participants to repeat digits in forward order and reverse order. Forward digit span is related to attention while reverse digit span is related to working memory. Performance is based on number of correct recitations.

The Grip Strength Test- (Reitan & Wolfson, 1985) is a test of hand strength in which the participant holds a dynamometer in the palm of his/her hand and squeezes the stirrup with his/her fingers at maximum intensity. This is repeated for the other hand, and the dynamometer indicates the participant’s score.
The Minnesota Multiphasic Personality Inventory – Restructured Format- (MMPI-2-RF; Ben-Porath & Tellegen, 2008) assesses major symptoms of psychopathology, personality characteristics, and behavioral proclivities.
Appendix B: Statistical Analyses

Outlier Analyses

Without statistical outliers, observed and unobserved controls did not perform differently on verbal memory tasks, $A = .96, F (4, 34) = .364, p = .832$, but performed differently on visual memory tasks, $A = .72, F (4, 34) = 3.36, p = .020$. With Bonferroni correction (.05/4), follow up ANOVAs indicated observed controls performed worse than unobserved controls on the CVMT d-prime task, $F(1,37) = 11.75, p < .0125, d = 1.1$, and the CVMT acquisition total $F(1,37) = 7.97, p < .0125, d = .88$. The groups did not perform differently on CFT recall, $F(1,37) = .35, p = .56, d = .19$; or CVMT delayed recognition, $F(1,37) = 1.63, p = .21, d = .42$.

The significant effect of non-credible effort on verbal memory also remained after removing the outliers, $V = .42, F (4, 82) = 14.87, p < .001$. Non-credible responders performed worse on RAVLT immediate recall $F(1,85) = 50.98, p < .0125, d = 1.60$; RAVLT 30-minute recall, $F(1,85) = 55.76, p < .0125, d = 1.67$; RAVLT Learning Over Trials, $F(1,85) = 16.56, p < .0125, d = .88$, and RAVLT 30-minute recognition, $F(1,85) = 49.17, p < .0125, d = 1.58$.

Absent outliers, non-credible participants performed worse than control participants on visual memory measures, $V = .43, F (4, 81) = 15.46, p < .001$. Non-credible participants performed worse on the CFT recall, $F(1,84) = 43.20, p < .0125, d = 1.44$; the CVMT d-prime task, $F(1,84) = 37.18, p < .0125, d = 1.35$; the CVMT acquisition total, $F(1,84) = 42.10, p < .0125, d = 1.46$; and the CVMT delayed recognition, $F(1,84) = 48.12, p < .0125, d = 1.52$. 
Chi Square Using WMT CNS

There was no difference between observed and unobserved non-credible
performance on WMT CNS failure, $\chi^2(1)=.52, p=.47$; 88.5% of observed non-credible
responders and 81% of unobserved non-credible responders failed the WMT CNS.

Counterbalanced Task Order Analyses

Order of task administration did not differentially effect observed or unobserved
control performance on verbal memory tests, $A = .918, F(4, 44) = .979, p=.429$ or visual
memory tests, $A = .894, F(4, 44) = 1.30, p=.286$. Non-credible participants’ and
controls’ performance on verbal memory tests and visual memory tests was not
differentially effected by task administration order, $V = .03, F(4, 92) = .585, p=.674; A = .942, F(4, 91) = 1.39, p=.243$. Order of task administration also did not differentially
effect observed or unobserved non-credible participant performance on verbal memory
tests, $A = .906, F(4, 41) = 1.06, p=.389$ or visual memory tests, $A = .934, F(4, 40) = .707, p=.592$.

Observation Beliefs

The proportion of participants who believed they were being videotaped was not
evenly distributed between observed and unobserved controls $\chi^2(1)=26.57, p =.000$ or
observed and unobserved non-credible responders, $\chi^2(1)=11.64, p =.001$. In all, 73.9%
of observed controls, 64% of observed non-credible responders, 3.7% of unobserved
controls, and 14.3% of unobserved non-credible responders believed they were being
videotaped.

After removing individuals who did not believe they were videotaped when in the
presence of a video camera, or did believe they were videotaped in the absence of the
video camera, the results of the hypothesis testing did not change. Observed and
unobserved controls did not perform differently on verbal memory tasks, $A = .969, F(4, 38) = .299, p = .877$. Observed controls performed worse than unobserved controls on visual memory tasks, $A = .717, F(4, 38) = 3.75, p = .011$. Using Bonferroni correction (.05/4), follow-up ANOVAs revealed observed controls performed worse than unobserved controls on the CVMT $d$-prime task, $F(1,41) = 12.66, p = .001$, and the CVMT acquisition total $F(1,41) = 12.73, p = .001$. The groups did not perform differently on CFT recall, $F(1,41) = .185, p = .67$, or CVMT delayed recognition, $F(1,41) = 1.35, p = .25$. There was no difference between observed and unobserved control participants on WMT IR failure, $\chi^2(1) = 2.25, p = .13$; 9.0% of observed controls and 0% of unobserved controls failed the WMT IR.

Collapsed across observation (including only individuals who believed they were observed), there remained a significant effect of non-credible effort on verbal memory measure performance, $V = .43, F(4, 72) = 13.74, p = .000$. Follow-up ANOVAs indicated non-credible responders performed worse on RAVLT immediate recall $F(1,75) = 42.56, p = .000$; RAVLT 30-minute recall, $F(1,75) = 44.74, p = .000$; RAVLT Learning Over Trials, $F(1,75) = 16.24, p = .000$ and RAVLT 30-minute recognition, $F(1,75) = 50.56, p = .000$.

When only participants who believed they were observed were included, collapsed across observation, non-credible participants performed worse than control participants on visual memory measures, $V = .36, F(4, 71) = 9.79, p = .000$. Follow-up ANOVAs indicated non-credible participants performed worse on the CFT recall,
$F(1,74)=22.73, \ p=.000$; the CVMT d-prime task, $F(1,74)=19.77, \ p=.000$; the CVMT acquisition total, $F(1,74)=22.15, \ p=.000$; and the CVMT delayed recognition, $F(1,74)=31.71, \ p=.000$. The proportion of WMT IR index failure was not evenly distributed between groups, $\chi^2(1)=40.62, \ p=.000$. In all, 80\% non-credible responders failed the WMT IR and 2.9\% control participants failed the WMT IR.

Observed non-credible participants and unobserved non-credible participants did not perform differently on the verbal memory measures, $\Lambda = .93$, $F(4, 29) = .52, \ p = .720$ or the visual memory measures, $\Lambda = .84$, $F(4, 28) = 1.38, \ p = .267$. There was no difference between observed and unobserved non-credible performance on WMT IR failure, $\chi^2(1)=.03, \ p = .86$; 78.6\% of observed non-credible responders and 81.3\% of unobserved non-credible responders failed the WMT IR.

**Compliance with Instructions**

A one-way ANOVA was utilized to test for differences in perceived success in ability to follow instructions, between groups. The four groups differed in how they rated perceived success following instructions, $F(3,95)=18.43, \ p=.000$. Post-hoc comparisons of the four groups indicated unobserved controls ($M = 4.54, 95\% \ CI [4.21, 4.87]$) rated their success with adhering to instructions significantly higher than observed non-credible responders ($M = 3.12, 95\% \ CI [2.77, 3.46]), $p = .000$ and unobserved non-credible responders ($M = 3.23, 95\% \ CI [2.86, 3.60]), $p = .000$. Observed controls ($M = 4.39, 95\% \ CI [4.03, 4.75]$) also rated their success with adhering to instructions significantly higher than observed non-credible responders ($M = 3.12, 95\% \ CI [2.77, 3.46]), $p = .000$ and unobserved non-credible responders ($M = 3.23, 95\% \ CI [2.86, 3.60]), $p = .000$. Observed
and unobserved control participants did not rate their perceived success with following instructions differently, $p = .560$; nor did observed and unobserved non-credible responders rate their perceived success with following instructions differently, $p = .660$.

**Control Participants’ Difficulty Ratings**

Collapsed across observation, control participants rated the CVMT delayed recognition, CFT delayed recall, and AVLT 30-minute delayed recall differently in terms of difficulty, $F (2, 78) = 11.97$, $p = .000$. Pairwise comparisons indicated the AVLT 30-minute recall was rated as more difficult than the CFT Recall, $p = .000$, yet as equally difficult as the CVMT delayed recall, $p = .419$. The CFT recall was rated as less difficult than the CVMT delayed recall, $p = .000$. There was no between groups effect, $F (1, 39) = .823$, $p = .370$.

**Table 5**

**Descriptive Statistics for Controls’ Difficulty Ratings for Delayed Recall Measures**

<table>
<thead>
<tr>
<th>Group</th>
<th>$N$</th>
<th>CFT Recall</th>
<th>CVMT DRT</th>
<th>RAVLT 30-min Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed Control</td>
<td>17</td>
<td>1.65 (.93)</td>
<td>2.59 (1.37)</td>
<td>2.59 (1.12)</td>
</tr>
<tr>
<td>Unobserved Control</td>
<td>24</td>
<td>1.50 (.83)</td>
<td>2.54 (1.18)</td>
<td>2.17 (.92)</td>
</tr>
</tbody>
</table>

*Note. Scores are included in mean (S.D.) format. Participants were asked to rate how difficult overall was the task, 5 being the most difficult. The test acronyms are as follows: Complex Figure Test (CFT), Continuous Visual Memory Test (CVMT), and Rey Auditory Verbal Learning Test (RAVLT).*

Across observation, control participants also rated the WMT IR, AVLT IR and CVMT acquisition total differently in terms of difficulty; using the Greenhouse-Geisser
correction, \( F (1.357, 57.02) = 7.20, p=.005 \). Pairwise comparisons indicated the WMT IR was rated as less difficult than the AVLT IR, \( p = .002 \) and the CVMT acquisition total, \( p = .000 \). The AVLT IR and CVMT acquisition total were not rated differentially difficult, \( p = .444 \). There was a significant between groups effect, \( F(1,42)=4.40, p=.042 \), indicating that overall, the observed control participants rated all tasks as more difficult than the unobserved controls. As the interaction approached significance, \( F (1.357, 57.02) = 2.35, p=.121 \), we decided to follow up on the interaction just to explore the nature of the interaction. Follow-up t-tests revealed, that observed controls did not rate WMT IR as more difficult than unobserved controls did, \( t(26.07)=1.87, p=.072 \).

Observed controls also rated the AVLT IR as equally difficult as unobserved controls, \( t(42)=-.427, p=.67 \). However, observed controls rated the CVMT acquisition total as more difficult than unobserved controls, \( t(42)=2.42, p=.02 \).

Table 6

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>WMT IR</th>
<th>CVMT Total</th>
<th>RAVLT IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed Control</td>
<td>20</td>
<td>1.85 (1.04)</td>
<td>3.0 (1.26)</td>
<td>2.2 (1.51)</td>
</tr>
<tr>
<td>Unobserved Control</td>
<td>24</td>
<td>1.38 (.49)</td>
<td>2.08 (1.25)</td>
<td>2.38(1.21)</td>
</tr>
</tbody>
</table>

*Note. Scores are included in mean (S.D.) format. Participants were asked to rate how difficult overall was the task, 5 being the most difficult. The test acronyms are as follows: Complex Figure Test (CFT), Continuous Visual Memory Test (CVMT), and Rey Auditory Verbal Learning Test (RAVLT).*
Non-Credible Participants Difficulty Ratings

Observed and unobserved non-credible participants rated feigning cognitive impairment equally difficult pre and post battery, $V = .11, F (2, 41) = 2.49, p = .10$. Across observation, non-credible participants also rated the WMT IR, CVMT acquaintance total and AVLT IR differently in terms of difficulty; $F (2, 76) = 8.25, p = .001$. Pairwise comparisons indicated the WMT IR was rated less difficult than the AVLT IR, $p = .010$ and the CVMT acquisition total, $p = .000$. The AVLT IR and CVMT acquisition total were not rated differentially difficult, $p = .369$. There was no between groups effect, $F (1, 38) = .037, p = .848$.

Table 7

Descriptive Statistics for Non-Credible Participants Difficulty Ratings for Immediate Recall Measures

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>WMT IR</th>
<th>CVMT Total</th>
<th>RAVLT IR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed Non-Credible</td>
<td>23</td>
<td>2.26 (1.36)</td>
<td>3.30 (1.33)</td>
<td>3.26 (1.25)</td>
</tr>
<tr>
<td>Unobserved Non-Credible</td>
<td>17</td>
<td>2.24 (1.39)</td>
<td>3.41 (1.62)</td>
<td>3.0 (1.66)</td>
</tr>
</tbody>
</table>

Note. Scores are included in mean (S.D.) format. Participants were asked to rate how difficult overall was the task, 5 being the most difficult. The test acronyms are as follows: Complex Figure Test (CFT), Continuous Visual Memory Test (CVMT), and Rey Auditory Verbal Learning Test (RAVLT).

Collapsed across observation, non-credible participants rated the CVMT delayed recognition, CFT delayed recall, and AVLT 30-minute delayed recall differently in terms of difficulty, $F (2, 62) = 6.83, p = .002$. Pairwise comparisons indicated the
AVLT 30-minute recall was rated less difficult than the CFT Recall, $p = .002$ and the CVMT delayed recognition, $p = .024$. The CFT recall was rated as difficult as the CVMT delayed recall, $p = .239$. There was no between groups effect, $F (1, 31) = 1.00$, $p = .324$.

Table 8

*Descriptive Statistics for Non-Credible Participants’ Difficulty Ratings for Delayed Recall Measures*

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>WMT IR</th>
<th>CVMT Total</th>
<th>CFT Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed Non-Credible</td>
<td>20</td>
<td>2.50 (1.15)</td>
<td>3.30 (1.22)</td>
<td>3.30 (1.13)</td>
</tr>
<tr>
<td>Unobserved Non-Credible</td>
<td>13</td>
<td>2.15 (.90)</td>
<td>2.77(1.24)</td>
<td>3.31(1.55)</td>
</tr>
</tbody>
</table>

*Note.* Scores are included in mean (S.D.) format. Participants were asked to rate how difficult overall was the task, 5 being the most difficult. The test acronyms are as follows: Complex Figure Test (CFT), Continuous Visual Memory Test (CVMT), and Rey Auditory Verbal Learning Test (RAVLT).
Appendix C: Other Documentation

Control Instructions

Title of Research: The effects of testing environment on cognitive performance
Researchers: Caitlin Reese, M.A. & Julie A. Suhr, Ph.D., Ohio University
Department: Psychology

Participant ID : ####

Today you will take a series of neuropsychological tests that assess motor speed, attention, memory, and thinking skills. You are asked to give your best effort on all of these tests. You will be taken to the testing environment in which you will take the tests in a moment.

Please do not share your instructions with anyone else, including the individual who gives you your tests, as different people have different instructions. If you have any questions, you may independently ask the individual who handed you these instructions.

Thanks for your participation.
Non-Credible Responding Instructions

Title of Research: The effects of testing environment on cognitive performance
Researchers: Caitlin Reese, M.A. & Julie A. Suhr, Ph.D., Ohio University
Department: Psychology

Participant ID : ####

Today you will take a series of neuropsychological tests that assess motor speed, attention, memory, and thinking skills. You were selected to participate in this study because you reported having a history of mild head injury on the online screen. Recall, if possible/applicable, what types of symptoms you temporarily experienced after your mild head injury. If you did not experience any symptoms, what types of symptoms do you think, in general, are associated with mild head injury? Now consider what would happen if those symptoms were significantly worse and permanent (e.g. representative of brain damage). You are being asked to believably pretend that you have significant problems (e.g. representative of brain damage) with motor speed, attention, memory, and thinking skills tests. In other words, pretend that you are someone involved in a lawsuit, and you want to pretend to have brain damage in order to win a financial settlement. What might you do to indicate (even though you do not) that you have permanent and significant problems with motor speed, attention, memory, and thinking skills while taking neuropsychological tests? There is no wrong answer. Of course, you would not want to be caught pretending, and therefore it is important to remind you to believably pretend to have such problems however you see fit. You will be taken to the testing environment in which you will take the tests in a moment.

Please do not share your instructions with anyone else, including the individual who gives you your tests, as different people have different instructions. If you have any questions, you may independently ask the individual who handed you these instructions.

Thanks for your participation.
Unobserved Consent Form

Title of Research: The effects of testing environment on cognitive performance
Researchers: Caitlin Reese, M.A. & Julie A. Suhr, Ph.D., Ohio University
Department: Psychology

You are being asked to participate in research. For you to be able to decide whether you want to participate in this project, you should understand what the project is about, as well as the possible risks and benefits in order to make an informed decision. This process is known as informed consent. This form describes the purpose, procedures, possible benefits, and risks. It also explains how your personal information will be used and protected. Once you have read this form and your questions about the study are answered, you will be asked to sign it. This will allow your participation in this study. You should receive a copy of this document to take with you.

Explanation of Study
You are invited to participate in a research study exploring performance on cognitive tests in different testing environments. While the study presents different testing environments, each participant will complete the cognitive tests in only one of those environments. The oral, computerized, and written tests assess cognitive abilities such as attention, verbal fluency, and general thinking skills. Participation is voluntary and may be discontinued at any time without penalty.

All tests will be administered by a trained graduate student who is supervised by the study director. We expect that your participation will take approximately two hours.

Risks and Discomforts
This study does not pose any known risks to you.

Benefits
Participants in this study will benefit from it by learning more about neuropsychological testing. Your participation will also enhance our understanding of what environmental situations are better suited for the most valid neuropsychological testing. Such knowledge would be important for neuropsychologists who practice in both clinical and forensic settings.

Confidentiality and Records
All information obtained from you in this study will be kept strictly confidential. This information will be identified according to a randomized, unique 4-digit subject identification number. There is no way that your name can or will be linked to your data. In addition, the data from this study will be kept in a locked storage facility and accessible only to authorized individuals.
Additionally, while every effort will be made to keep your study-related information confidential, there may be circumstances where this information must be shared with:
  * Federal agencies, for example the Office of Human Research Protections, whose responsibility is to protect human subjects in research;
  * Representatives of Ohio University (OU), including the Institutional Review Board, a committee that oversees the research at OU;

**Compensation**

You will receive two experimental credit points for participation in this study. If for any reason you feel you need to discontinue your participation in this study, you will receive one experimental credit point reflecting the time you spent in the study (e.g. If you discontinue within the first hour of the study, you will receive one point. If you discontinue within the second hour you will receive two points).

**Contact Information**

If you have any questions about your participation, please do not hesitate to ask the experimenter. You may also contact the study co-investigators, Caitlin Reese, M.A. or Julie A. Suhr, Ph.D. if you have additional questions or concerns.

If you have any questions regarding your rights as a research participant, please contact Jo Ellen Sherow, Director of Research Compliance, Ohio University, (740)593-0664.

By signing below, you are agreeing that:
- you have read this consent form (or it has been read to you) and have been given the opportunity to ask questions
- known risks to you have been explained to your satisfaction.
- you understand Ohio University has no policy or plan to pay for any injuries you might receive as a result of participating in this research protocol
- you are 18 years of age or older
- your participation in this research is given voluntarily
- you may change your mind and stop participation at any time without penalty or loss of any benefits to which you may otherwise be entitled.

Signature __________________________________________ Date ____________

Printed Name __________________________________________

Version Date: [12/10/07]
Observed Consent Form

Title of Research: The effects of testing environment on cognitive performance
Researchers: Caitlin Reese, M.A. & Julie A. Suhr, Ph.D., Ohio University
Department: Psychology

You are being asked to participate in research. For you to be able to decide whether you want to participate in this project, you should understand what the project is about, as well as the possible risks and benefits in order to make an informed decision. This process is known as informed consent. This form describes the purpose, procedures, possible benefits, and risks. It also explains how your personal information will be used and protected. Once you have read this form and your questions about the study are answered, you will be asked to sign it. This will allow your participation in this study. You should receive a copy of this document to take with you.

Explanation of Study
You are invited to participate in a research study exploring performance on cognitive tests in different testing environments. While the study presents different testing environments, each participant will complete the cognitive tests in only one of those environments. You may or may not be videotaped. The oral, computerized, and written tests assess cognitive abilities such as attention, verbal fluency, and general thinking skills. Participation is voluntary and may be discontinued at any time without penalty.

All tests will be administered by a trained graduate student who is supervised by the study director. We expect that your participation will take approximately two hours.

Risks and Discomforts
This study does not pose any known risks to you.

Benefits
Participants in this study will benefit from it by learning more about neuropsychological testing. Your participation will also enhance our understanding of what environmental situations are better suited for the most valid neuropsychological testing. Such knowledge would be important for neuropsychologists who practice in both clinical and forensic settings.

Confidentiality and Records
All information obtained from you in this study will be kept strictly confidential. This information will be identified according to a randomized, unique 4-digit subject identification number. There is no way that your name can or will be linked to your data. In addition, the data from this study will be kept in a locked storage facility and accessible only to authorized individuals.
Additionally, while every effort will be made to keep your study-related information confidential, there may be circumstances where this information must be shared with:
* Federal agencies, for example the Office of Human Research Protections, whose responsibility is to protect human subjects in research;
* Representatives of Ohio University (OU), including the Institutional Review Board, a committee that oversees the research at OU;

**Compensation**

You will receive two experimental credit points for participation in this study. If for any reason you feel you need to discontinue your participation in this study, you will receive one experimental credit point reflecting the time you spent in the study (e.g. If you discontinue within the first hour of the study, you will receive one point. If you discontinue within the second hour you will receive two points).

**Contact Information**

If you have any questions about your participation, please do not hesitate to ask the experimenter. You may also contact the study co-investigators, Caitlin Reese, M.A. or Julie A. Suhr, Ph.D. if you have additional questions or concerns.

If you have any questions regarding your rights as a research participant, please contact Jo Ellen Sherow, Director of Research Compliance, Ohio University, (740)593-0664.

By signing below, you are agreeing that:
- you have read this consent form (or it has been read to you) and have been given the opportunity to ask questions
- known risks to you have been explained to your satisfaction.
- you understand Ohio University has no policy or plan to pay for any injuries you might receive as a result of participating in this research protocol
- you are 18 years of age or older
- your participation in this research is given voluntarily
- you may change your mind and stop participation at any time without penalty or loss of any benefits to which you may otherwise be entitled.

Signature_____________________________________________ Date ____________

Printed Name__________________________________________

Version Date: [12/10/07]
Demographics Questionnaire

DIRECTIONS: Please choose the best response for each question.

1. What is your age?
   A. 18
   B. 19
   C. 20
   D. 21
   E. 22
   F. 23
   G. 24
   H. 25
   I. Other (please specify) ________________

2. What is your current year in college?
   A. First
   B. Second
   C. Third
   D. Fourth
   E. Fifth or above
   F. Graduate student
   G. Other

3. What is your current overall college GPA?
   A. less than or equal to 2.5 (less than C)
   B. 2.51 to less than 3.0 (C to B)
   C. 3.0 or greater (B or better)

4. What is your racial/ethnic identity?
   A. Caucasian, Non-Hispanic
   B. African American
   C. Latino or Hispanic
   D. Asian or Pacific Islander
   E. American Indian or Alaska Native
   F. Two or more races
   G. Other

5. Are you currently diagnosed with depression, anxiety, or other psychological condition?
   A. No
   B. Yes
If yes, what condition(s)____________________________

6. Are you currently receiving treatment (e.g., prescription medication, counseling, herbal supplements) for depression, anxiety, or another psychological condition?
   A. No
   B. Yes
If yes, what treatment(s)?___________________________

7. To the best of your knowledge, do you have a diagnosed learning disorder?
   A. No
   B. Yes
If yes, which one(s)___________________________

8. To the best of your knowledge, are you currently diagnosed with any neurological deficit, impairment, or diagnosis?
   A. No
   B. Yes
If so, which one?_______________________________
Debriefing Form

Title of Research: The effects of testing environment on cognitive performance
Researchers: Caitlin Reese, M.A. & Julie A. Suhr, Ph.D., Ohio University
Department: Psychology

Thank you for your participation. The study you have just participated in examined whether videotaping people would change their performance on cognitive tests. Furthermore, the study examined whether the presence of a video camera would affect the performance of those participants asked to malingering (i.e., pretend to have brain damage) while completing the tasks. **People who pretend to have brain damage (malingerers) usually do so because they want to win a lawsuit or similar incentive, and they often do so after they have sustained only a mild head injury. In order for our results to generalize, or apply, to this population, participants (including yourself) were recruited because they reported having experienced a mild head injury in the past.**

Some participants were asked to do their best on the cognitive tests, and some participants were asked to pretend to have brain dysfunction while taking the tests. Those individuals who were asked to pretend to have brain dysfunction received web-based information about head injury; this information was accurate yet incomplete. **The website from which the head injury information came was a lawyer’s website that over-exaggerates and generalizes head injury symptom severity.** It is true that groups of people who have sustained head injuries sometimes perform worse on cognitive tests than people who have not. However, the groups who tend to perform poorly have much more severe head injuries than average. **Furthermore, many individuals who sustain a mild head injury do not have persistent cognitive deficits following the injury.** For more information regarding this fact, please see the following journal articles:


Given your high level of functioning (attending college), it is unlikely that you have any persistent cognitive deficits related to mild head injury. However, if you have additional questions, please ask your test administrator and/or contact Dr. Julie Suhr (contact information below).

Some participants’ consent forms indicated that they may or may not be videotaped, that their performance would only be accessible to authorized researchers, and that all video tapes would be discarded at the end of the quarter. These individuals also saw video cameras in their test rooms and were told their evaluations would be video recorded in session. Other participants did not receive any of these instructions, and did not have video cameras in their rooms. While some participants were told their evaluations would be video-recorded, this statement was for the purposes of the study, and NO ONE’s evaluations were video recorded.

Some prior research suggests that the presence of a third party observer, even when the observer is a video camera, can have detrimental effects on cognitive performance, but how it will affect individuals asked to simulate poor performance is not yet known. The data participants anonymously provided will be used to determine whether significant relationships validate these hypotheses.

As many other people will be participating in this study during the quarter, we ask that you please do not share details about this study with anyone else, so that people remain unaware of the specific details of the study given above.

If you have additional questions or concerns regarding your participation, please contact one of the co-investigators below.

Graduate Researcher:  Caitlin Reese, M.A.
044O Porter Hall
cr145708@ohio.edu.

Faculty Advisor:  Julie Suhr, Ph.D.
241 Porter Hall
suhr@ohio.edu

In addition, if you are concerned about the study materials used or questions asked and wish to speak with a professional, or if you would like more information or reading material on this topic, please contact one of the following resources:

Ohio University Counseling and Psychological Services: 593-1616

Ohio University Psychology and Social Work Clinic: 593-1092
If you have any questions regarding your rights as a research participant, please contact Jo Ellen Sherow, Director of Research Compliance, Ohio University, (740)593-0664.