1-1-2013

Effects of third party observation behind a one-way mirror on neuropsychological tests with varying conative load

Jessica Stenclik
University at Albany, State University of New York, jhstenclik@gmail.com

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Part of the Psychology Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive.
Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
EFFECTS OF THIRD PARTY OBSERVATION BEHIND A ONE-WAY MIRROR ON NEUROPSYCHOLOGICAL TESTS WITH VARYING CONATIVE LOAD

by

Jessica H. Stenclik

A Dissertation

Submitted to the University at Albany, State University of New York

in Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

College of Arts & Sciences

Department of Psychology

2013
Effects of Third Party Observation Behind a One-Way Mirror on Neuropsychological Tests with Varying Conative Load

by

Jessica H. Stenclik

COPYRIGHT 2013
Acknowledgements

It is with immense gratitude that I acknowledge various individuals who in one way or another contributed and extended their invaluable support during the completion of this project. I wish to thank, first and foremost, my mentors Dr. Robert J. McCaffrey and Dr. Julie K. Lynch for their continued guidance and assistance throughout the preparation and completion of this project. They have facilitated my academic growth and inspired me. I am also incredibly thankful to Dr. Paul Green for his generosity in providing necessary resources for the completion of this dissertation. I would further like to thank my husband Derek Stenclik for his unselfish and unfailing encouragement and inspiration, without which all of this would not be possible. Lastly, I am eternally grateful to my loving and supportive family for their constant understanding and reinforcement throughout my academic and life journey.
# Table of Contents

Chapter 1 ................................................................................................................................. 1

The Use of Third Party Observers ......................................................................................... 1

Issues Regarding Third Party Observation During Neuropsychological Assessment ...... 6

Validity and Reliability .............................................................................................................. 6

Distraction ................................................................................................................................. 7

Standardization and Normative Data ...................................................................................... 8

Test Security and Misuse ......................................................................................................... 9

Confidentiality .......................................................................................................................... 11

Recommendations for Responses to Requests for TPOs for Neuropsychologists .......... 11

Social Facilitation ..................................................................................................................... 15

Arousal Theories ...................................................................................................................... 16

Activation/Drive Theory .......................................................................................................... 16

Learning Theory/Evaluation Apprehension ............................................................................. 16

Monitoring Theory .................................................................................................................. 18

Attentional Conflict/Information Processing Theories ......................................................... 19

Distraction-Conflict Theory .................................................................................................... 19

Attention Process Model ........................................................................................................ 19

Automatized Processing and Attention ................................................................................ 20

Challenge and Threat Explanation ......................................................................................... 20

Social Valuation Theories ....................................................................................................... 22

Objective Self-Awareness Theory .......................................................................................... 22

Control Systems Model .......................................................................................................... 22
Self-Presentational Account ......................................................... 23

Valence Anticipate Evaluation/Self-Efficacy Theory ........................... 23

Social Facilitation Effects on Experimental Tasks ............................... 24

Motor .......................................................................................... 25

Visuo-Motor .................................................................................. 27

Memory ......................................................................................... 27

Mental and Visual Tracking .............................................................. 29

Abstraction and Concept Attainment ............................................... 30

Language and Comprehension ......................................................... 30

Signal Detection and Attention ....................................................... 31

Miscellaneous ............................................................................... 31

Videotaping, Audiotaping, Computer Monitoring and the use of One-Way Mirrors ..... 33

Third Party Observers and Social Facilitation Effects on Neuropsychological Tests ..... 35

Conation, Third Party Observation, Task Performance ........................ 40

Purpose ....................................................................................... 43

Chapter 2 ..................................................................................... 43

Method ......................................................................................... 43

Participants ................................................................................... 43

Material ........................................................................................ 45

Finger Tapping Test ....................................................................... 45

Trail-making Test A and B ............................................................... 46

Category Test ............................................................................... 47

The California Verbal Learning Test-II .............................................. 48
Medical Symptom Validity Test ................................................................. 48
Digit Span from the Wechsler Adult Intelligence Scale- 3rd edition ................. 49
Procedure .................................................................................................... 49
Overview of Data Collection and Analysis .................................................. 50
Chapter 3 .................................................................................................... 51
Results ........................................................................................................ 51
Demographics ............................................................................................ 51
Performance on Neuropsychological Measures that Differ in Conative Load .... 51
MANCOVA ............................................................................................... 51
Univariate ANVOCA s ............................................................................... 53
Performance on the CVLT-II ...................................................................... 54
MANCOVA ............................................................................................... 54
Univariate ANCOVA s ............................................................................... 57
Chapter 4 .................................................................................................... 58
Discussion ................................................................................................... 58
Key Findings .............................................................................................. 60
Clinical Implications .................................................................................. 65
Limitations and Conclusions ..................................................................... 67
References .................................................................................................. 70
Table 1: Rank-Order of 19 Tests from the HRNB in Order from Lowest to Highest Conative Load ................................................................. 90
Table 2: Criteria for Failure on Measures of Performance Validity .................. 91
Table 3: Demographic Information for Total Group, Examiner Only Condition, and TPO Condition ................................................................. 92
Table 4: Means and Standard Deviations for Five DVs Differing Conative Load Included in the First MANCOVA for the Examiner Only Condition and TPO Condition .......... 93

Table 5: Means and Standard Deviations on the Five DVs Differing in Conative Load Organized by Gender ........................................................... 94

Table 6: Means and Standard Deviations for Six DVs from the CVLT-II Included in the Second MANCOVA for the Examiner Only Condition and TPO Condition ................. 95

Appendix A: Instructions for the Trail-making Test ......................................................... 96

Appendix B: Category Test Instructions ................................................................. 100

Appendix C: Instructions for each condition read to the examinee ........................... 102
Abstract
This study investigated the effects of third party observation with a one-way mirror on tests that differ in conative load from the Halstead Reitan Neuropsychological Battery (HRNB). Additionally, the California Verbal Learning Test- II (CVLT-II) was examined, as this is a popular measure of memory with which the effects of third party observation is unknown. Examinees (n=93) were recruited from an undergraduate research pool of a large State University. All were administered the Medical Symptom Validity Test (MSVT), Finger Tapping Test (FTT), Digit Span from the Wechsler Adult Intelligence Scale- 3rd edition, Trail Making Test A and B, CVLT-II, and the Category Test (CT). Participants were randomized to two conditions: 48 were observed by a third party behind a one-way mirror and 45 had only the examiner present. Performance validity was assessed using the MSVT, Reliable Digit Span, and CVLT-II forced choice recognition. All participants were identified as putting forth valid performance. Analyses revealed no significant differences in performance on any the measures between the two conditions. Gender was found to have a significant effect on scores on FTT-dominant hand, FTT-nondominant hand, and Trails B from the HRNB, and List B from the CVLT-II. This study found no effects of third party observation with the use of a one-way mirror on tests of varying conative load from the HRNB or on six indices of the CVLT-II. Implications for forensic practice are addressed.

Keywords: third party observer, conation, Halstead Reitan Neuropsychological Battery, California Verbal Learning Test-II
Chapter 1

The Use of Third Party Observers

The presence of a third party observer (TPO) during a neuropsychological exam has given rise to strongly positioned discussions, considerations, and recommendations. A TPO is an individual separate from the examinee and examiner who observes and possibly documents a neuropsychological or psychological evaluation (Otto & Krauss, 2009). An observer can be any third party, however, more often than not they are parents of the examinee, trainees for educational purposes, supervisors, translators, other neuropsychologists or psychologists, attorneys, legal assistants, or another type of health care professional (McSweeny, Becker, Naugle, Snow, Binder, & Thompson, 1998). Third party observation can also be completed indirectly, without the presence of another individual, through the use of recording devices such as audio and video recorders, as well as one-way mirrors. Historically, TPOs have been most frequently present in a medical setting during the assessment and feedback of medical procedures such as imaging and diagnostic testing. Although these methods are relatively resistant to TPO effects, the nature of neuropsychological and psychological evaluations differs greatly from medical procedures. Neuropsychological and psychological examinations have a greater potential of being influenced and/or altered by the presence of a TPO due to the importance of relational variables implicit throughout the evaluation such as rapport, as well as the importance of concentration and valid performance (McCaffrey, Fisher, Gold, & Lynch, 1996).
In a position paper addressing the effects of third party observation on neuropsychological examinations and recommendations for the use of TPOs, the American Academy of Clinical Neuropsychology (AACN; 2001) differentiates between two different types of TPOs: (1) an involved TPO and (2) an uninvolved TPO. An uninvolved TPO is considered an observer who is impartial to the outcome of an examination. These TPOs are often times those in training or supervisory positions who are most interested in learning the process and skills necessary for a neuropsychological evaluation. They may also be interested in the behavior of the examinee and examiner as well as a particular clinical diagnosis (AACN, 2001). Third party facilitators are a subtype of uninvolved TPOs who are utilized to aid in the examination process and include translators and parents of children with problems that may inhibit the evaluation process such as high levels of anxiety or behavioral disturbances (Otto & Krauss, 2009; McSweeny, 1998). Either way, an uninvolved TPO holds no interest in the specific patient in particular or the outcome of the examination. It is important note, however, that an uninvolved TPO may still have a significant effect on functioning of both the examiner and the examinee due to their presence in the examination room (Duff & Fisher, 2005). Alternatively, an involved TPO holds a stake in the outcome and consequences of the neuropsychological examination. For example, an involved TPO may be interested in the outcome of the evaluation in regards to legal, financial, relational, and/or social implications (AACN, 2001).

TPOs can occur in either a clinical/medical or forensic/medicolegal context. TPOs may be commonly used in the clinical evaluation of children; however they are most frequently utilized within a forensic setting regarding both criminal and civil cases.
(Axelrod et al., 2000; Duff & Fisher, 2005; McCaffrey, 2005). These can include issues surrounding personal injury litigation, competency evaluations, not guilty by reason of insanity pleas, and workers compensation and/or disability claims (Cramer & Brodoskly, 2007). Forensic evaluations, such as an independent medical evaluation, are quite different from clinical evaluations because the neuropsychologist is not hired by the patient, but rather a third party. Therefore, the neuropsychologist is responsible to the third party and does not work for the patient (Bush & NAN Policy & Planning Committee, 2005). As a result, the relationship between the examiner and patient, as well as the responsibilities that the neuropsychologist has to the patient, is limited. For example, a doctor-patient relationship does not exist and confidentiality is limited because the examiner is reporting to the third party. Regardless of the presence of a doctor-patient relationship, it is nonetheless essential that the neuropsychologist proceed in an ethical and professional manner (Bush et al., 2005).

The right to a TPO and legal representation during a criminal forensic evaluation is non-negotiable. Alternatively, TPOs are not permitted to be present during psychiatric, psychological, and/or neuropsychological examinations in the federal court. Civil forensic matters, however, are less clear and frequently TPOs are requested to be present during these evaluations (AACN, 2001). Most states are unclear regarding the allowance of TPOs in the examination room during a civil matter and instead consider the demand for a TPO and requests to prohibit observation by a TPO on an individual basis (McCaffrey et al., 1996). For example, although North Carolina gives psychologist the right to deny TPOs, in some jurisdictions it is an examinees right to have their independent medical examination observed (Bush et al., 2005; Duff & McCaffrey, 2005).
Therefore, when TPOs are demanded during a forensic evaluation, conflict arises between the ethical responsibilities of a neuropsychologist, the legal rights of the examinee, as well as the law (Duff & McCaffrey, 2005; Howe & McCaffrey, 2012).

When examined on a case-by-case basis, a lawyer may present multiple arguments for the allowance of a TPO during the evaluation. One particular argument may be that of transparency. Lawyers believe that the use of a TPO is beneficial to the patient in that it provides an opportunity for inspection and analysis of the methods and procedures used in order to assist in questioning (Howe & McCaffrey, 2012). Lawyers may hold the belief that observing an evaluation will aid in their ability to challenge methods or results that are believed to be inaccurate at trial due to their ability to record any missteps made throughout the evaluation (Shuman, 1994). Observation is also thought to ensure honesty and fairness, the administration of valid assessment instruments and procedures, as well as strict adherence to test protocols (Otto & Krauss, 2009; McCaffrey et al., 1996; McCaffrey & Howe, 2012; Simon, 1996). Furthermore, attorneys may want to be present to monitor the discussion between the examiner and examinee and provide live consultation to their client to avoid damaging statements and to guarantee that the neuropsychologist does not question the plaintiff inappropriately (AACN, 2001; Simon, 1996). Shealy, Cramer, & Pirelli (2008) also discuss the issue that defendants themselves may be hesitant or deny participation in a neuropsychological evaluation without their attorney present.

Benevolent attorney motives have also been cited as reasons for observing a neuropsychological assessment. For example, being present provides the opportunity to meet and become familiar with the examiner providing the assessment as well as may
help facilitate and encourage discussion and communication between the examiner and examinee. However, neuropsychologists argue that these interactions can be carried out through other means rather than the use of TPOs and are not contingent on participating as a TPO. For example, attorneys have the ability to consult with other forensic neuropsychologists and can utilize deposition to question the examiner (Cramer & Brodosky, 2007; Simon, 1996).

Attorneys may also be prone to cite both the Fifth and Sixth Amendments to the United States Constitution as grounds to include a TPO during a forensic evaluation (Shealy et al., 2008). The Sixth Amendment discusses rights related to criminal prosecutions and includes the right that a criminal should have the assistance of counsel for his/her defense, while the Fifth Amendment protects against the abuse of government authority in a legal procedure. This is likely one reason that TPOs are allowed during criminal forensic evaluations, however this issue is not as clear regarding civil forensic cases. It is important to note, though that in 1981 the U.S. Supreme Court was in agreement with a lower court’s ruling that attendance of an attorney during a forensic evaluation would not only have little benefit but may also be harmful to the nature of the evaluation (Estelle v. Smith). Contrary to this, in New York an appellate court decided during Parsons v. Hytech Tool and Die, Inc. that a court inappropriately required a plaintiff have a neuropsychological evaluation without a TPO present. Therefore, it is evident that both state and federal courts are not in agreement regarding the presence of TPOs during neuropsychological examinations (Shealy et al., 2008).

An anonymous questionnaire sent out to 250 randomly selected members of NAN in 2001 found that the majority of neuropsychologists (88.0%) have never had an
attorney sit in during neuropsychological testing (Mittenberg, Petersen, Strauman, & Cooper, 2001). Furthermore, more than half (53.0%) reported they have never had an attorney ask to be present during the evaluation. These results were interpreted as evidence that neuropsychologists successfully decline attorneys the right to sit in on neuropsychological examinations with their clients (Essig, Mittenberg, Petersen, Strauman, & Cooper, 2001). In contrast, in 2008, a survey of 400 randomly sampled licensed psychologist who were members of Division 41 (American Psychology-Law Society) of the American Psychological Association found that approximately three-fourths of psychologists conducting criminal forensic evaluations have had a TPO present during an examination such as interpreters, students or trainees, and defense attorneys. However 76.0% never allowed either video or audio recording of the evaluation (Shealy et al., 2008). In addition, the majority of psychologists believed that the presence of a TPO likely has negative effects on the results of the evaluation by affecting the validity of the examination and confidentiality (Shealy et al., 2008). Given the conflicting results between the two surveys, it is clear that not all clinicians are in agreement regarding the presence of attorneys during an evaluation and the possible implications that a TPO can have on a neuropsychological examination.

**Issues Regarding Third Party Observation During Neuropsychological Assessment**

**Validity and reliability.** The fundamental issue central to the argument against the presence of TPOs during a neuropsychological examination revolves around the validity and reliability of the assessment results (Axelrod et al., 2000). It is likely that the presence of a TPO alters rapport between the examiner and examinee which may make the examinee feel uncomfortable, decrease the amount of information that is obtained by
the examiner, and influences the examinees motivation level and behavior (AACN, 2001; Duff & Fisher, 2005). The Committee on Psychological Tests and Assessments of the American Psychological Association (2007) state that not only are test results likely skewed, but also interview and observation may not provide valid information in the presence of a TPO because examinees are less likely to be open with the examiner and share sensitive or personal information. The presence of a TPO may not only change the behavior of the examinee, but also affect the examiner. For example, the examiner may feel anxious or nervous and act differently in the presence of an attorney or TPO and be prone to making mistakes in test administration thereby decreasing the validity of the assessment results (Duff & Fisher, 2005; Otto & Krauss, 2009). Thus, it is proposed that the inclusion of a TPO during a neuropsychological evaluation leads to an invalid and inaccurate assessment and limits the utility and applicability of test results and interpretations associated with decision making (Duff & Fisher, 2005).

**Distraction.** The Standards for Educational and Psychological Testing state that the environment within which testing occurs should contain minimal distractions and be reasonably comfortable (American Educational Research Association (AERA), 1999). Lezak (2012) suggests that the validity of a neuropsychological assessment is called into question due to decrements in performance as a result of an adverse and distracting testing environment. Therefore, it is essential that the testing environment be tightly controlled and free from distractions in order to ensure the validity of the assessment (Antastasi, 1988; Lynch & McCaffrey, 2004). Observation by a third party likely increases the risk of distraction and may also lead to decreased comfort on both the part of the examiner and examinee due to increases in pressure.
AACN (2001) discusses the difference between two types of distractions that can result from third party observation: (1) external distraction and (2) internal distraction. External distraction is defined as observable stimuli such as sneezing/coughing, body movement, and the sound of notes being taken or paper rustling. These types of distractions can be decreased by removal of the TPO from the testing environment through the use of indirect observation with audio or video recording and one-way mirrors. However, internal distractions are imperceptible, reside within the examinee or examiner, and likely lead to feelings of social discomfort and distraction due to concerns of evaluation. Internal distractions are not removed through the use of recording or a mirror because the perception of being observed is still present. The effects of distraction are not only a concern during forensic evaluations, but also clinical evaluations in which supervisors, trainees, parents, or translators are acting as the TPO (AACN, 2001).

**Standardization and normative data.** The American Psychological Association Ethical Standards (2002) 9.02 (Use of Assessments) and 9.06 (Interpreting Assessment Results) encourage the use of appropriate assessment measures with adequate reliability and validity, in addition to providing accurate interpretation consistent with the assessment tool utilized. When limitations to the evaluation results exist (e.g., presence of a TPO), these should be stated in the report in the section discussing the interpretations. Daubert criteria from Daubert v. Merril Dow Pharmaceuticals (1993) also mandate that valid and reliable assessment methods be used as a basis for expert opinion. Furthermore, Standard 5.1 of the Standards for Educational and Psychological Testing instructs those administering assessment measures to strictly adhere to the standardized administration and scoring procedures in the test manual (AERA, 1999).
Neuropsychological and other psychological tests are standardized in a controlled environment without the presence of a TPO (Axelrod et al., 2000; McCaffrey et al., 1996). Additionally, standardized administration of these assessment measures likely does not include the presence of a TPO. In addition, if the examinee has had a prior evaluation in which a TPO was not present, comparison of results from an exam that was administered in the presence of a TPO is unreliable due to difference in administration methods (McSweeny et al., 1998). Therefore, the presence of a TPO during test administration and the departure from standardized practice is not only in contrast to the APA Ethical Standards, AERA Standards, and Daubert Criteria, but more importantly makes comparison to normative data and prior assessments inappropriate and invalid (Howe & McCaffrey, 2012; Lynch & McCaffrey, 2004; McCaffrey et al., 1996). Without the appropriate test administration and comparison to normative data, the interpretation of test results are limited and provide less utility at both the clinical and forensic level.

**Test security and misuse.** Psychologists and neuropsychologists also have a responsibility to ensure that assessment measures are secure and utilized appropriately. It is incredibly important that examinees presenting for neuropsychological tests are unaware of the test questions and items they are going to be asked to complete as prior knowledge of the measures creates an unfair advantage that will render results invalid (Axelrod et al., 2000; Howe & McCaffrey, 2012). Psychologists are also ethically obligated to maintain the security of test materials through the APA Ethical Standard 9.11 and Standard for Education and Psychological Testing Standard 5.7 (AERA, 1999; APA, 2002).
The presence of a TPO increases the risk of a test security breach, as neuropsychologists are unable to control how a TPO will utilize the information about the assessment measures they obtain while observing an evaluation (AACN, 2001; McCaffrey et al., 1996). For example, if the TPO is a lawyer or legal paraprofessional, they may be tempted to utilize their knowledge of the assessment instruments to coach future litigants undergoing an evaluation so that the client appears to present with a specific neurological or psychiatric condition (Howe & McCaffrey, 2012; McCaffrey et al., 1996). Therefore, the presence of a TPO may decrease the utility and integrity of neuropsychological assessment instruments and violate ethical standards to maintain test security (McCaffrey et al., 1996).

In addition to test security, the ethical standards of the profession also address issues regarding test misuse. Ethical Standard 1.01 (Misuse of Psychologist’s Work) addresses the importance of ensuring that a psychologist’s work is neither misused nor misrepresented. In addition, Ethical Standard 9.07 (Assessment by Unqualified Persons) discourages the use of assessment measures by unqualified persons unless it is for training or educational purposes (APA, 2001). Once again, the psychologist has no control over how a TPO will utilize their familiarity with the assessment measures and this could result in various types of test misuse including unqualified test administration. Duff and Fisher (2005) also discuss potential copyright issues that may arise due to the presence of a TPO. For example, psychologists are responsible to uphold copyright laws of the tests they are purchasing. Reproduction of this material by a TPO is considered a violation of the purchasing agreement. Thus, it is evident that observation by a TPO may
place the neuropsychologist at risk for violating several ethical standards as well as copyright laws due to the misuse and lack of test security in the presence of a TPO.

**Confidentiality.** Complications regarding confidentiality are especially salient in the context of clinical evaluations that may involve a friend, relative, or significant other as a TPO. Ethical Standard 4.01 discusses the psychologist’s responsibility to maintain confidentiality and take measures to ensure information obtained from the patient or examinee remains confidential (APA, 2002). When a TPO is present in a clinical setting, it is the psychologists’ ethical responsibility to ensure that the patient is aware of the limits to confidentiality and the possibility that the TPO may disclose information from the evaluation to others. In the setting of a forensic evaluation, confidentiality is already limited due to the lack of a doctor-patient relationship. Even within the context of already limited confidentiality, however, a TPO still has the ability to disclose sensitive information to uninvolved parties who hold no authorization to the information (Shealy et al., 2008). The presence of a TPO, therefore, decreases the probability of the right to confidentiality the examinee is warranted, regardless of context.

**Recommendations for Responses to Requests for TPOs for Neuropsychologists**

Some researchers in the field of neuropsychology believe that a neuropsychologist should do everything in their power to deny a request for a TPO during a forensic evaluation (Axelrod et al., 2000). McCaffrey (2005) and Yantz and McCaffrey (2005) state that it would be prudent to bar trainees and supervisors as TPOs for educational purposes in cases with serious consequences, such as competency evaluations and various types of litigation. Others believe that it is appropriate to include the TPO during the
clinical interview portion of a neuropsychological evaluation; however, TPOs should be removed during the testing segment of the assessment (McSweeny et al., 1998). In addition, it has been stated that alternate means of observation such as audio or video recorders and/or a one-way mirror that have similar implications as the physical presence of a TPO, may be inappropriate, and should also be excluded (Axerlrod et al., 2000; AACN, 2001; NAN, 2000). Contrary, Cramer and Brodsky (2007) and the Committee on Psychological Testing and Assessment (2007) believe that a videotaped examination or observation behind a one-way mirror may be appropriate as long as it is out of sight of the examinee and not easily noticeable.

Professional organizations have also put forth statements regarding the presence of TPOs during an evaluation. As stated above, AACN (2001) states that TPOs should be permitted to be present during criminal cases but prohibited from observation in the majority of civil cases. There may be civil cases in which it is appropriate to include a TPO, though, for example during an evaluation of a child or older adult with behavioral difficulties. The presence of a parent in this example may actually allow testing to move forward and encourage the child to engage in the assessment process (AACN, 2001; Axelrod et al., 2000). Additionally, the National Academy of Neuropsychology (NAN) put forth a consensus paper indicating that examiners should employ every effort to avoid the presence of a TPO during an evaluation; however, recognize that there are some circumstances in which it may be appropriate to include a TPO, such as for training and supervisory purposes or when evaluating children (NAN, 2000). Furthermore, the American Psychological Association’s Committee on Psychological Tests and Assessment (2007) state that a neuropsychologist should recommend the withdrawal of a
request for a TPO or minimize the impact the TPO has on the observation if a TPO will be present. In addition, this committee also notes that it is important the neuropsychologist is aware of their right to refuse to perform the assessment in the presence of a TPO.

These organizations have also addressed the use of alternative or secretive recording devices such as one-way mirrors, audio recording, and video recording rather than the actual physical presence of a third party observer (Bush, Pimental, Ruff, Iverson, Barth, & Broshek, 2009). The American Psychological Association’s Committee on Psychological Tests and Assessment (2007) takes the stance that in the context of an evaluation in which observation must occur, psychologists should utilize methods that interfere the least, such as with a one way mirror. AACN (2001) takes a very different stance on this issue and argues that the use of these indirect means of observation should be considered inappropriate as they have the same effect on an examinee as the physical presence of a third party observer. This group proposes that the presence of a recording device or one-way mirror provides the same internal distraction as the physical presence of a TPO and does not decrease other issues, such as test security, associated with a TPO discussed above. Additionally, the use of these methods without informing the examinee has also been suggested as a way to avoid possible TPO effects (Bush et al., 2009). NAN has taken a stance on this issue, however, arguing that it is unethical to deceive the examinee by withholding information that they are being observed or recorded. Furthermore, secretive recording can also alter the examiner’s behavior and does not ensure test security (Bush et al., 2009).
Many recommendations have been made regarding the best way to address requests for TPOs. The APA Ethical Standard 1.02 states that when there is a conflict between ethics and the law, it is the psychologist’s responsibility to attempt to resolve the conflict (APA, 2002). This indicates that due to the ethical issues associated with the presence of a TPO, a neuropsychologist presented with a request for a TPO in a forensic context should make attempts to educate the legal community (e.g., judge or attorney requesting the TPO) as well as the examinee in order to bring to light implications a TPO may have on the validity of the evaluation (Lynch & McCaffrey, 2004; McCaffrey, 1996). In order to do so, the neuropsychologist should be well versed in ethical obligations and standards, current research regarding the effect of TPOs and social facilitation, position papers and policy statements from professional organizations, and state statutes (Howe & McCaffrey, 2012; McCaffrey et al., 1996).

Although it is recommended to exclude the presence of TPOs to the best of the neuropsychologists’ abilities, there are times when a TPO is either court ordered or the neuropsychologist agrees to include a TPO during an evaluation. Given these circumstances, there are multiple steps that should be taken to ensure that the TPO has a minimal impact on test results. First, the neuropsychologist should attempt a compromise with the TPO and allow observation during the initial interview but disallow observation during the testing process (Duff & Fisher, 2005; McSweeny et al., 1998). If the TPO will be observing the formal testing procedures, the neuropsychologist should utilize all strategies to ensure that the TPO has as minimal impact as possible. This will maximize the validity of the results. For example, the TPO should sit behind the examinee and be instructed that they are not allowed to speak during the exam (Committee on
Psychological Tests and Assessment, 2007). In addition, it has been proposed that when a TPO is present it would be best to use assessment instruments that have been found to be less affected by social facilitation as well as those that provide increased protection of the test stimuli, such as self-report measures, to protect test security (Committee on Psychological Tests and Assessments, 2007; Otto & Krauss, 2009). To further promote test security and confidentiality when a TPO is mandated, the examiner should advocate that the TPO be an individual ethically obligated to protect test security and confidentiality such as a psychologist, receive an agreement from the TPO that they will uphold the test security and confidentiality, as well as attempt to obtain a protective order in a forensic context (Howe & McCaffrey, 2012; Otto & Krauss, 2009). Lastly, documentation of the TPO’s presence is essential following an evaluation. Written reports should document the effects that the TPO may have had on the validity and interpretation of the interview, observations, and test results (Duff & Fisher, 2005; Howe & McCaffrey, 2012).

Social Facilitation

Theories concerning the influence of a third party observer on neuropsychological test performance are rooted in the social psychology literature. The phenomenon, referred to as social facilitation, is a concept dating back as far as the late nineteenth century when Triplet (1898) uncovered that cyclists rode faster when they were racing in groups as compared to racing alone. In general, social facilitation is the tendency for individuals to perform better on simple tasks (e.g., more accurately or more quickly) but worse on complex tasks (e.g., less accurately or slower) in the presence of others. The fundamental and most popular social psychology theories concerning social facilitation
and the effects of the mere presence of an examiner as well as the effect of evaluation during task performance are discussed below.

**Arousal theories.**

*Activation/Drive Theory.* Zajonc (1965) based his Drive Theory hypothesis off of the Hull-Spence position (Spence, 1956). He postulated that an increase in drive would lead to an increase in well-learned and dominant responses and a decrease in new, less well-learned responses. Zajonc (1965;1966) proposed that an increase in arousal and activation as a consequence of being observed results from the unpredictability inherent in the situation. The presence of another person is believed to increase the individual’s arousal or drive level and therefore lead to an increase in dominant and/or easy responses and a decrease in subordinate and/or difficult responses. Zajonc (1965) believed that the mere physical presence of an observer, even if they are not present as an evaluator, will result in increases in arousal and hence an increase in drive. This innate increase in drive is a reflexive response that results in more accurate simple, well-learned behaviors but inaccurate novel or difficult responses.

*Learning Theory/Evaluation Apprehension Model.* Cottrell, Wack, Sekerak, and Rittle (1968) extended Zajonc’s (1965) Drive Theory by hypothesizing that it is not only the mere presence of an individual that results in social facilitation. Rather, they believed that the examinee expects evaluation from the presence of an audience and this evaluation apprehensions results in increased drive and arousal. Cottrell, Rittle, and Wack (1967) found that the presence of an audience improved performance on a simple list learning task, yet impaired performance on a more difficult task, validating Zanjoc’s
(1965) Drive Theory; however, Cottrell, et al. (1968) found that the mere physical presence of an audience did not have the same effect. Only when the audience was a spectator and engaged in observing the task did a drive effect occur. Therefore, Cottrell et al. (1968) speculated that anticipation of both positive and negative evaluations from others observing performance leads to increased arousal and drive and a social facilitation effect. Henchy and Glass (1968) also found support for the evaluation apprehension effect. Participants who believed they were being evaluated compared to those in the presence of a non-evaluating observer or alone had greater frequencies of dominant compared to subordinate responses.

Sanders (1981) utilized a classical conditioning model to further elaborate Cottrell et al.’s (1968) Learning Theory. Sanders (1981) noted that individuals are classically conditioned to expect either reward or punishment following evaluation by others. Therefore, social facilitation results whenever evaluation occurs due to the association of certain consequences with evaluation. Thus, any type of signal (including but not limited to the mere presence of others) should increase drive and arousal as a result of the expectations associated with the evaluation by others leading to a social facilitation effect (Sanders, 1981).

Zajonc (1980) further extended his original Drive Theory as a response to Cottrell et al.’s (1968) and Henchy and Glass’ (1968) theories including the effects of participant apprehension of evaluation with his Alertness Theory. This theory states that drive and arousal are increased as a result of the participant’s uncertainty as to how the observer is going to behave. Zajonc (1980) postulated that it is likely that evaluation apprehension effects are present as a result of this unpredictability; however, an effect due to the mere
presence of an audience will still occur. He referred to this mere presence effect as a compresence effect that will occur regardless of the influence of other, more significant effects (i.e., evaluation apprehension).

**Monitoring Theory.** Guerin and Innes (1982) build upon prior arousal theories by adding the component of monitoring. Monitoring is defined as a “broad process of orienting and attending to stimuli to assess familiarity, threat and impending interactions or encounters” and occurs as a result of the mere presence of another individual (Guerin & Innes, 1982). Monitoring Theory states that increases in arousal will likely result in the presence of another individual when the audience’s behavior is unpredictable, difficult to monitor effectively and easily, or complex and difficult to interpret. This increase in arousal and distraction caused by the attempt and inability to monitor another individual effectively is due to the participant’s perception of physical threat and results in a social facilitation effect (Guerin & Innes, 1982). Guerin (1983) stated that alertness and arousal is mediated by increased monitoring due to the threat of an audience rather than just the mere presence of another individual.

In 1977, Geen and Gange reviewed twelve years of social facilitation research and concluded that Drive Theory (Zajonc, 1965) is the most parsimonious explanation for social facilitation effects. The authors did, however, recognize that Cottrell’s (1968) Evaluation Apprehension Model also provides a reasonable explanation. Geen and Gange (1977) stated at the end of their review that they believed future models would be more likely to utilize and integrate cognitive theories of social facilitation into the current activation theories.
Attentional conflict/ information processing theories.

**Distraction-Conflict Theory.** Sanders and Baron (1975) suggested that the relationship between observation and arousal is mediated by distraction. The presence of an audience in the room results in the division of attentional resources between the task at hand and the present individual. This attentional conflict results in increased arousal, which then leads to the social facilitation effect. Furthermore, both social and non-social stimuli have the ability to cause distraction. For this reason, the authors believed that Cottrell’s (1968) model is not sufficient to account for social facilitation effects due to it’s requirement of the presence of a social being to create evaluative apprehension. Thus, distraction acts as a drive factor and interacts with the task at hand (Sanders & Baron, 1975).

Baron, Moore, and Sanders (1978) described the Distraction-Conflict Theory as the drive that results from an attentional conflict between a task and audience distraction that will facilitate simple tasks yet impair complex tasks, leading to a social facilitation effect. They stated that when examinees are under observation they monitor themselves by searching for cues and signals from their audience. This monitoring results in increased distraction. Simple and well-learned tasks are facilitated by the distraction because examinees only need a central focus on simple tasks; however, complex and novel tasks are impaired due to decreased attention to peripheral information that may be necessary for task success (Baron et al., 1978).

**Attentional Process Model.** Sanders (1981) noted that facilitation only occurs if the increase in drive and arousal is greater than the examinees’ impaired attention and
further extended the Distraction-Conflict Theory into an Attentional Process Model. This model integrates principles from a mere presence model, learned drive model, and the Distraction-Conflict theory. Sanders (1981) stated that initially, an individual completing a task in the presence of an audience will engage in a reflexive learned drive and arousal response; however, maintenance of the arousal effects occur as a result of attentional conflict.

**Automatized Processing and Attention.** Manstead and Semin (1980) created a model that also utilizes cognitive factors as a mediator between the presence of an audience and the consequences of social facilitation. The Automatized Processing and Attention model considers the effects of an audience on the conscious monitoring of tasks. Specifically, Manstead and Semin state that for well-learned, routine, easy tasks, individuals are less likely to monitor their performance when working alone, resulting in suboptimal performance. With the addition of an audience, examinees concentrate more intently on the task at hand and this consequently leads to improved performance. The gains made by observation are due to a shift from automatic information processing to greater controlled processing. Dissimilarly, examinees are more likely to attend at higher levels to complex and novel tasks rather than well-learned tasks without the presence of an audience. Thus, the presence of an audience acts as a distracting factor to the examinee and places greater demands on the examinee’s attention, resulting in impaired performance (Manstead & Semin, 1980).

**Challenge and Threat Explanation.** Blascovich, Mendes, Hunter, and Salmon (1999) considered the effect of both threat and challenge cognitive appraisals as a contributing moderating factor to the social facilitation effect. Challenge appraisals are
those that occur before a stressful event when individuals feel prepared and competent to cope with a stressor. In the context of a challenge appraisal, the individual will feel increased drive and motivation to perform and will anticipate success. Contrary, a threat appraisal occurs when the individual feels the stressor outweighs the individual’s coping resources. This results in feelings of pessimism regarding task completion, and feelings of danger and doubt. Blascovich et al. (1999) proposed that challenge appraisals occur when individuals are faced with a well-learned task with an audience present, while threat appraisals will be prevalent if the individual must complete a challenging, complex task in the presence of an audience. Hence, challenge appraisals during simple tasks results in more accurate performance and threat appraisals during complex tasks will lead to a deficit in performance. Furthermore, individuals who make challenge compared to threat appraisals were found to each have unique physiological responses (Blascovich et al., 1999).

More recently, Feinberg and Aiello (2012) supported Blascovich et al.’s (1999) claims and concluded that cognitive appraisals are an important moderating element between the presence of an audience and task performance. Specifically, examinees who received “challenge instructions” performed better than examinees administered “threat instructions” in the presence of others. Alternatively, when alone, those in the threat condition showed better performance than those in the challenge condition. Interestingly, these findings were consistent across tasks of varying difficulty and did not support an interaction between task difficulty and the presence of others. Thus, Feinberg and Aiello concluded that the way in which an individual appraises a task in the presence of an audience is a core factor in the individual’s future performance.
Social valuation theories.

**Objective Self-Awareness Theory.** Objective Self-Awareness Theory takes into consideration aversive cognitive states that arise due to increases in an individual’s objective self-awareness and decreases in subjective self-awareness as a result of observation (Wickland & Duval, 1971). Subjective self-awareness takes place when an individual attends to and evaluates other individuals. On the other hand, objective self-awareness occurs when an individual focuses their attention inward and evaluates their own performance. During objective self-awareness an individual’s self-consciousness increases and they will compare themselves to either their own personal standards or those of society. Wickland and Duval (1971) theorized that in the presence of an audience, an individual’s objective self-awareness will increase causing the individual to attempt to perform consistently with his/her or societies standards. This will then lead to evaluation and recognition of the discrepancy between his/her performance and the standards that they set for themselves resulting in an aversive cognitive state. These negative cognitions increase drive and motivation on simple tasks resulting in improved performance, but this increased effort exceeds one’s capabilities and results in deficits in performance on more complex tasks (Wicklund & Duval). Wickland and Duval find it important to note, however, that the presence of others is not necessary for this model. The only essential component is the reduction of self-awareness and any stimuli that prompts the individual to attend to themselves and their own performance.

**Control Systems Model.** Similar to Wicklund and Duval (1971), Carver and Scheier (1980) also discussed the effect of increased self-attention on social facilitation and the importance of the comparison an individual makes between their own
performance and a selected standard of behavior. Dissimilarly, Carver and Scheier argued that it is not an aversive cognitive state that facilitates simple tasks, but rather a behavioral feedback loop. This feedback mechanism is a discrepancy-reducing feedback loop in which an audience increases examinee self-attention resulting in an increased awareness of the discrepancy between one’s performance and the social standard leading to task performance that conforms more closely to social standards (Carver & Scheier). This feedback loop is thought to facilitate performance on simple tasks; however, no explanation is provided for performance on complex tasks.

**Self-Presentational Account.** Bond (1982) discussed the importance an individual places on managing their self-image and the effect this may have on task performance in the presence of another individual. More specifically, when in the presence of others, an individual is aware of discrepancies between their actual and model performance and increase motivation and their drive to reduce any inconsistencies. This then results in enhanced performance on simple tasks and diminished performance on complex tasks. Bond also emphasized the emotion of embarrassment in his model. Specifically, performance will only be influenced when an individual experiences embarrassment due to incorrect responding or poor performance.

**Valence Anticipated Evaluation and Self-Efficacy Theory.** Sanna and Shotland (1990) proposed that the valence of the evaluation an individual expects to receive from an audience influences the type of effect (good or bad) an audience will have on performance. Individuals will expect a positive evaluation from others if they anticipate performing well on a task; however, they will expect a negative evaluation if they believe they will perform poorly. Intuitively, individuals expect to perform better on simple
tasks than complex tasks. Thus, individuals will show improved performance in front of an audience if they have high expectations for success due to increased drive and strategies to preserve their self-presentation, which is more likely to occur during simple tasks. Dissimilarly, more complex tasks will result in low expectations for success and impairment in performance.

Providing support for Sanna and Shotland’s (1990) Valence Anticipated Evaluation model, Sanna (1992) developed the Self-Efficacy Theory. This theory postulates that an individual’s drive is established by their efficacy and outcome expectancy. Efficacy expectancy occurs when an individual believes they have the skills necessary for success, while outcome expectancy is the belief that particular behavior will ensure a specific outcome. If an individual has high efficacy and outcome expectancy in the presence of an audience, most common on simple tasks, their performance will be enhanced. Conversely, when an individual has low efficacy and outcome expectancies in the presence of an audience (i.e., when performing complex tasks) impaired performance results (Sanna).

Social Facilitation Effects on Experimental Tasks

The social psychological literature provides extensive examples of research originating in the late nineteenth century offering support for and against the effects of social facilitation on various tasks (Aiello & Douthitt, 2001). Variations in performance as a result of social facilitation and the presence of an audience have been examined on numerous tasks including those assessing motor, visuo-motor, visual tracking, memory,
abstraction, language and comprehension, signal detection and attention, and miscellaneous performance. Specific examples are notes below.

**Motor.** The effect of social facilitation on motor tasks has been found to vary regarding task type and task complexity. Multiple studies have found a positive relationship between audience observation and task performance. For example, Baker, Jung, and Petrella (2011) found that muscular strength in the form of bench and leg presses increases significantly for males and females in the presence of an observer. Physical activity such as jogging is also improved in the presence of an audience and runners tend to accelerate when being observed (Stube, Miles, & Finch, 1981; Worringham & Messick, 1983). Social facilitation effects have also been evidenced on tests of physical fitness such as number of sit-ups and times on a shuttle run with children with mental retardation (Bowman & Dun, 1982). Furthermore, simple motor tasks such as pulling a lever have been found to increase in the presence of an audience (Fouts, 1979).

Evidence also exists that the effects of social facilitation on motor task varies depending on task complexity. Hass and Roberts (1975) utilized complex tracing task in the presence of an audience. Participants who underwent a learning trial prior to the task were facilitated by the presence of an audience, while participants who were learning the motor task for the first time during the evaluation process showed inhibited performance. Additionally, Bell and Lee (1989) evaluated the effects of an audience on both the number and accuracy of karate kicks for participants who were skilled versus unskilled in karate. The presence of an audience had no effect on the accuracy of kicks for skilled participants; however, resulted in a decreased frequency of accurate kicks for unskilled
participants. Thus, it is evident that prior knowledge and learning of a task may determine whether the effects of social facilitation are present.

Dissimilarly, motor performance has been found to be unaffected or impaired in the presence of an audience. In contrast to Fouts (1979), Clark and Fouts (1973) examined lever pulling performance in the presence of either a negative or positive audience. Clark and Fouts (1973) found significantly lower lever pulling intensity when a positive audience observed participants, a finding inconsistent with their hypothesis that participants anticipating positive outcomes will show improved performance due to increases in motivation. The authors suggest that these findings may be explained by the idea that participants in the presence of a positive audience were less anxious to perform, had less drive, and therefore did not show improvements in performance. Sorce and Fouts (1973) also had participants engage in a lever pulling task, except dissimilarly from the above studies, differentiated between those with high, medium, and low motivation. Findings revealed that participants high in motivation performed worse in the presence of an audience than those with low or medium motivation (Sorce & Fouts).

Additionally, although Hass and Roberts (1975) found that social facilitation effects on a tracing task were dependent on task complexity, Innes and Young (1975) utilized one tracing task and found inhibited performance in the presence of an audience. This may indicate that Innes and Young’s (1975) tracing task was complex rather than simple. Lastly, Sasfy and Okun (1974) discovered that motor performance on a “roll-up” task was impaired when individuals were evaluated both directly and indirectly by those considered experts, while no effects were found in a non-expert condition indicating that
evaluation may play an important role in the effects of social facilitation on motor performance.

**Visuo-motor.** The majority of research examining social facilitation on visuo-motor performance has found that the presence of an audience inhibits task performance. Specifically, Lombardo and Catalano (1975) examined performance on a rotary-pursuit task and found that the presence of an expert audience decreases participant performance. Lombardo and Catalano examined performance on this task for participants who first failed a task in the presence of an audience compared to those who did not perform a first task. Participants who failed a first task and performed the rotary-pursuit task in the presence of an audience performed the poorest (Lombardo & Catalano). Gore and Taylor (1973) provide further evidence for decrements in performance on a rotary pursuit visuo-motor task due to the presence of an expert audience, while dissimilarly, Guerin (1986) found that participants actually performed better on a simple rotary-pursuit task in the presence of an audience. Lastly, Miller, Hurkman, Feinberg, and Robinson (1979) found differences based on task complexity on this type of task. The presence of an audience impaired performance on difficult rotary pursuit tasks yet facilitated performance on an easy rotary pursuit task. On the other hand, Innes and Gordon (1985) found no interaction between task difficulty and presence of an audience and overall found more errors on both a difficult and simple eye-hand coordination task in the presence of an audience.

**Memory.** Multiple researchers have examined the effects of an audience on both immediate and delayed recall utilizing a paired-associates learning task. Studies have consistently found that in the presence of an observer, participants asked to immediately
recall items from the list (usually after two minutes) perform worse than participants who are alone. Alternatively, delayed recall, approximately 45 minutes after list-learning has occurred is facilitated in the presence of an audience compared to in isolation (Deffenbacher, Platt, & Williams, 1974; Geen, 1971; Geen, 1973).

Rather than manipulating the time before recall (i.e., immediate versus delayed), other researchers have explored the difference between hard (e.g., use of competitive list) and easy (use of non-competitive list) list learning in the presence of an audience. These studies have found, similarly to the studies above, that the presence of an observer results in facilitative effects on recall of the easy list, but impaired recall of the difficult list (Berkley & Hope, 1972; Feinberg & Aiello, 2006; Geen, 1983). The number and status of audience members has also been found to influence the effects of an observer on paired-associates learning. Specifically, Seta, Seta, Crisson, and Wang (1989) discovered that participants will experience increased feelings of anxiety as well as an increase in errors when observed by an audience member of a high status compared to one high and one low status observer indicating that the presence of a low status observer may decrease the levels of anxiety felt through observation. Lastly, Huston (1970) had participants learn two paired-associate lists: one learned alone and one in the participants of an audience. Results indicate that when an observer is present during the learning of a second list following the learning of a first list alone, retroactive inhibition was reduced and participants were able to recall more words from the first list.

Other memory tasks have also been utilized to examine social facilitation effects. Crisson, Seta, and Seta (1995) utilized a task in which participants were asked to memorize twenty-three digit numbers under specific categories denoted by color.
Crisson et al. (1995) measured participants’ expectancies in an attempt to determine if their efficacy expectancies, as manipulated by receiving positive or negative feedback on practice tasks, influenced memory performance in the presence of an audience. It was found that low-expectation participants showed impaired performance in front of an audience rather than alone; however, high-expectancy participants had facilitated performance in the presence of an audience. Therefore, expectations for success and failure may moderate the effects of social observation on memory performance (Crisson et al., 1995). Additionally, Seta, Seta, Donaldson, and Wang (1988) gave participants a list of words to categorize either alone or when being observed behind a one-way mirror. Participants recalled an equal number of words in both conditions, however, were less successful in organizing information when being observed. These findings may indicate that observation results in reduced amounts of mental capacities available for organizational processing (Seta et al., 1988). Lastly, Hanawalt and Ruttiger (1944) found that the presence of an audience facilitates story recall by increasing the number of facts and words that are remembered. These studies provide support for social facilitation and inhibition effects on various memory tasks.

**Mental and visual tracking.** Social facilitation effects on mental and visual tracking has been examined most often with the use of maze tasks. For example, Abel (1938; 1939) was the first to evidence social facilitation effects on maze tracing in the presence of an observer for two groups of individuals with suboptimal intelligence levels: (1) those with intelligence quotients (IQs) between 50-59 and (2) those with IQs in the range of 70-79. Rajecki, Ickes, Corcoran, and Lenerz (1977) also found that participants traced a maze more quickly and accurately in the presence of an audience. Task
complexity, however, has also been found to moderate the effects of an observer on task performance. Particularly, task performance was facilitated on simple maze tasks and impaired on complex maze tasks in the presence of an audience (Griffith, Fichman, & Moreland, 1989; Hunt & Hillery, 1973; Miyamoto, 1979; Shaver & Liebling, 1976). Liebling and Shaver (1973) also found support for a social inhibition/loafing effect on a complex maze when an observer was present, but no significant difference on performance of a simple task in the presence of an observer compared to alone.

**Abstraction and concept attainment.** Researchers have consistently found that the presence of an observer on both abstract anagram tasks requiring participants to rearrange letters to form words as well as concept attainment problems impairs performance relative to controls who perform these tasks in isolation (Aiello & Svec, 1993; Geen, 1977; Laughlin and Jaccard, 1975; Laughlin & Wong-McCarthy, 1975). Laughlin and Wong-McCarthy also found that impairment occurred not only in the physical presence of another individual but also when participants were videotaped, filmed, and audiotaped. Social loafing rather than facilitation effects are likely found on these tasks because of their complexity. Both anagram and concept attainment problems are difficult relative to other tasks and, consistent with theories of social facilitation, performance on these tasks is therefore impaired in the presence of an observer.

**Language and comprehension.** The effects of social facilitation on language construction and comprehension have also been assessed using experimental tasks in the social psychology literature. Rittle and Bernard (1973) engaged participants in a simple word construction task either in the presence of an experimenter with his back to the participant, with an experimenter observing the participant explicitly, or alone. The
authors found a social facilitation effect and increased response rates in both the mere presence and observation condition rather than alone (Rittle & Bernard, 1973). Listening comprehension through the use of an audio-taped informational speech about computer music has also been examined (Beatty, 1980). Listening comprehension may have increased complexity in comparison to word construction, as Beatty (1980) found decreased comprehension in the presence of an audience compared to in isolation.

**Signal detection and attention.** Studies have also examined participants’ abilities to discern between different stimuli and their ability to attend to different patterns using signal detection. Findings have been mixed. Although both Patania and Moran (2001) and Putz (1975) found support for social facilitation in the presence of an audience or one-way mirror during signal detection tasks, Kushnir and Duncan (1978) discovered that participants being observed actually made more errors when in the presence of an audience rather than alone. Attention has also been assessed using a vigilance task in which participants press a button when they see a dot flash on a television screen (Harkins, 1987). Harkins’ study also provides support for the effects of social facilitation on tasks of attention, in that participants whose output was evaluated performed better than those whose output was not evaluated.

**Miscellaneous.** Researchers examining the influence of observation on performance have utilized a variety of miscellaneous tasks as well, including computer tasks, video games, musical tasks, and unsolvable or unobtrusive tasks. Similar to research above, social facilitation effects vary based on task type, task difficulty, and participant characteristics. For example, individuals with high expectations for success performed better on a computer game than those with low expectations for success when
in the presence of an audience (Robinson-Stavely & Cooper, 1990). Computer tasks with varying difficulty also affect the impact of observation. Schmitt, Gilovich, Gore, and Joseph (1986) found that participants performed better on a simple computer task and worse on a complex computer task when in the presence of an audience.

Video game performance has also been shown to vary both by gender and task complexity when under observation. Specifically, males performed better on an electronic video game when in the presence of an observer, but this effect was not found for females (Ferris, 1985). Additionally, on a complex video game, participants performed worse in the presence of an audience than alone; however, no difference was found on a simple video game task (Kimble & Rezabek, 1992). Dissimilarly, on musical tasks, evaluation in person or through a one-way mirror had no effect in comparison to performance in isolation (Abrams & Manstead, 1981).

Performance has also been evaluated on an unsolvable task as well as an unobtrusive task with no performance criteria in place to decrease evaluation anxiety (Geen, 1981; Markus, 1978). On an insoluble task, participants who were informed they performed well on a first task performed better on a second task in the presence of an audience; however, if participants were told they performed poorly on a first task there were no differences in performance between those being observed compared to those alone, providing evidence for the evaluation apprehension hypothesis (Geen, 1981). Lastly, on a task of dressing and undressing in both familiar (i.e., simple task) and unfamiliar (i.e., complex task) clothing, participants showed facilitated performance while being observed on the easy task and inhibited performance when observed on the difficult task, consistent with social facilitation theories (Markus, 1978). Thus, it is
evident that the effects of social facilitation are dependent on different criteria for various as well as the factors manipulated within each study.

**Videotaping, Audiotaping, Computer Monitoring and the use of One-Way Mirrors**

The majority of these social psychology studies had participants perform various tasks either alone or in the physical presence of an examiner or observer. A number of studies, however, have also utilized alternative methods of observation that do not involve direct presence of another party but rather indirect monitoring of the participant’s performance. Different methods of indirect observation rather than direct physical observation have been used, with the most common modes being videotaping, audiotaping, and the use of one-way mirrors.

Video and audio taping have been used as methods of monitoring task performance on musical tasks, paired-associate memory tasks, signal-detection tasks, and letter-copying tasks (Abrams & Manstead, 1981; Geen, 1973; Putz, 1975; Terry & Kearnes, 1993), many of which are described above. In short, the effects of both video and audiotapes were found to have the same effect on performance as the direct evaluation conditions (Geen, 1973; Putz, 1975), including studies in which no social facilitation effects were found (Abrams & Manstead, 1981). These findings indicate that indirect observation through the use of audiotaping and/or videotaping have effects on performance similar to the presence of a physical audience.

Computer monitoring rather than direct physical observation has also been found to produce social facilitation effects. More specifically, Aiello and Svec (1993) discovered complex task performance on a computer anagram task to be impaired when
participants were monitored both by a physical audience as well as through a computer. Similarly, Douthitt and Aiello (2001) found complex task performance when solving problems on a computer was impaired when participants were monitored by a computer and they had no control over turning the off the monitoring. If participants had control over the monitoring and were able to switch it on and off, no difference in results was evident (Douthitt & Aiello, 2001). Moreover, computer monitoring during a simple data correction task improved task performance providing further evidence for social facilitation effects through the use of indirect computer monitoring (Stanton & Julian, 2002).

One-way mirrors or windows have also been utilized as ways to indirectly monitor and observe various types of task performance. For example, Criddle (1971) found that the physical presence of an audience is not necessary for social facilitation effects to occur on a paired-associates recall task. Performance on this task was impaired when participants were observed through a one-way mirror (Criddle, 1971). Social loafing, or impaired performance with indirect observation through a one-way mirror, has also been found on tasks assessing participants ability to learn a list of nonsense syllables as well as a task requiring participants to categorize and organize lists (Ganzer, 1968; Seta et al., 1988). Alternatively, social facilitation and improvements in performance have also been evidenced on a signal detection task as well as a simple noncompetitive paired associates task when observation occurs behind a one-way mirror (Musick, Beehr, & Gilmore, 1981; Putz, 1975).

Overall, findings from the social psychology literature indicate that the presence of others both directly and indirectly through the use of recording devices and one-way
mirrors most often results in impaired performance on complex tasks and facilitation on simple tasks. Specifically, observation often will lead to less accurate and slower performance on tasks of increased complexity and quicker performance and fewer errors on simplistic tasks (Bond & Titus, 1983). Bond and Titus completed a meta-analysis of 241 studies from 1898-1983 and stated that although these effects are present, social facilitation on simple tasks and social loafing on complex tasks account for a small amount of variance in experimental task performance. Specifically, approximately one to three percent of the variance accounted for in participant performance on complex tasks and only about 30 percent of the variance accounted for in performance on simple tasks is due to the presence of others (Bond & Titus).

Third Party Observers and Social Facilitation Effects on Neuropsychological Tests

Given the possible consequences associated with the use of TPOs during neuropsychological evaluations (McCaffrey et al., 2005) and the social psychology literature evidencing the effects of social facilitation on task performance (Bond & Titus, 1983), it is especially important to investigate social facilitation effects that a third party may have on neuropsychological test performance. It is significant to note that the social psychology literature cannot be generalized to inform the understanding of third party observers of neuropsychological testing. The majority of the studies from the social psychology literature have the participant performing the task either alone or in the presence of one observer (i.e., the examiner). Neuropsychological settings with TPOs are clearly different in that neuropsychological assessment always occurs in presence of an examiner. The question remains about the effects of an additional observer (i.e., the third party) on neuropsychological test performance (Eastvold, Belanger, & Vanderploeg,
Research examining the effects of a TPO on neuropsychological assessment instruments tapping various cognitive domains is discussed below.

Binder and Johnson-Greene (1995) reported the first case in which a TPO was found to affect neuropsychological performance. Utilizing a natural ABAB design, the authors found that an adult showed increased impairment on the Portland Digit Recognition Test, a forced-choice measure of symptom validity when her mother was present during testing. More current research has found that the presence of a third party causes decrements in performance on measures of executive functioning as tested with the Tactual Performance Test and Stroop Color and Word Test (Horowitz & McCaffrey, 2008; Huguet, Galvaing, Monteil, & Dumans, 1999; Kehrer, Sanchez, Habif, Rosenbaum, & Townes, 2012). Performance on tests of attention is also inhibited in the presence of a TPO on digit span tasks, the Stroop Color and Words Test, and Paced Auditory Serial Addition Test (Kehrer et al., 2012; Quarter & Marcus, 1971). Kehrer et al. also found performance on tasks of phonemic verbal fluency to be impaired in the presence of a TPO, while Wagstaff et al. (2008) found phonemic and semantic fluency task performance to be impaired due to a decreased number of switches, although there was no effect on word total. Lastly, Keher et al. discovered a performance decrement with verbal learning and memory as measured with the Rey Auditory Verbal Learning Test in the presence of a TPO.

The studies above utilized undergraduate students to examine the effects of a TPO on performance; however, negative effects of a TPO have also been evidenced with both children and clinical populations. Specifically Lynch (2005) examined 60 individuals who experienced a closed head injury. These individuals were given the Verbal Paired
Associates Tests from the Wechsler Memory Test-Revised, Trailmaking Test A and B, Finger Tapping Test, Grooved Pegboard, and Grip Strength with either only the examiner present or the examiner and a TPO present. Results indicated that delayed recall on the Verbal Paired Associates Test was impaired due to the presence of a TPO (Lynch, 2005). Additionally, Yantz and McCaffrey (2009) examined the effects of parental presence on alternate forms on test of nonverbal intelligence (the Test of Nonverbal Intelligence-3rd Edition) and memory (Selective Reminding Test) with children aged six to eight. An interaction between parental presence and order of observation was found. Particularly, children showed worse nonverbal intelligence when parents observed the first compared to second round of testing (Yantz & McCaffrey).

Variations in mode of observation have also been examined. Constantinou, Ashendorf, and McCaffrey (2002) examined the influence of audio recording on neuropsychological task performance. Performance on tasks of learning and memory (the Memory Assessment Scales (MAS)), verbal span (from the MAS), and motor functioning (Finger Tapping, Grooved Pegboard, and Grip Strength) were examined. Result indicate that when participants were knowledgeable that they were being audio-recorded, their performance on measures of learning, recall, and verbal span was poorer than those who were unaware of the recording device. No difference in performance was found on motor tasks (Constantinou et al., 2002). The effects of video recording on tasks of learning, recall, and verbal span from the MAS as well as on the same motor measures from the previous study have also been examined (Constantinou, Ashendorf, & McCaffrey, 2005). Performance was found to be impaired in the presence of video
recording on tasks of learning and recall, with no differences found on motor measures consistent with the study above (Constantinou et al., 2005).

More similar to the social psychology literature, Yantz and McCaffrey (2007) studied the influence of only examiner attention versus inattention on computerized tasks of symptom validity and executive functioning. A third party was not utilized. The examiner had either her back to the participant as they completed the tasks or watched the participant as they completed three tasks: the Test of Memory Malingering (TOMM), the Word Memory Test, and the Wisconsin Card Sorting Test. Examiner attention was found to have a facilitating effect on the TOMM. Specifically, participants who performed the TOMM on the computer with the examiner watching made fewer errors (Yantz & McCaffrey, 2007).

Position statements regarding the presence of a TPO suggest that TPOs are appropriate in training situations (AACN, 2001; Bush et al., 2009). Yantz and McCaffrey (2005) completed a study to determine if the presence of an examiner’s supervisor during neuropsychological testing impairs memory performance similar to the effects of other TPOs. In the observation condition, participants were made explicitly aware that the TPO was present only to evaluate the examiner’s performance, not the participant’s performance on the tests. Regardless of this information, the presence of the supervisor resulted in poorer performance on tests of memory (Yantz & McCaffrey, 2005). These findings indicate that irrespective of the reason for a TPO, the presence of a third party may still result in impaired test performance.
As an attempt to decrease the effects of a TPO on task performance, Gavett and McCaffrey (2007) employed the use of an adaptation period in order to allow participants to adjust to the presence of a TPO. Given that most studies do not find a TPO effect on motor functioning, motor measures were used to fill the adaptation period and familiarize the participant with performing tasks in front of both an examiner and TPO in thoughts that this would decrease TPO effects on later memory measures. Unfortunately, the adaptation period was not found to decrease TPO effects on memory performance and currently there are no empirically validated methods to avoid TPO effects on task performance (Gavett & McCaffrey).

One meta-analysis (Eastvold, Belanger, & Vanderploeg, 2012) and one summary study (Gavett, Lynch, & McCaffrey, 2005) examined past literature to inspect the effect and variance accounted for on neuropsychological task performance due to the presence of a third party observer. After reviewing 36 social psychology social facilitation studies as well as several neuropsychological studies, Gavett et al. found that an observer had an average effect size of .17 in the social facilitation literature and .10 when using only neuropsychological studies. Further, the presence of a TPO on cognitive tests assessing memory had a medium effect, while with tasks assessing attention, executive functioning, and motor functioning a TPO had a small average effect size (Gavett et al.).

Eastvold et al. (2012) criticized Gavett et al. (2005) because their study did not report whether the observer improved or inhibited performance as well as questioned Gavett et al.’s method for estimating the effect sizes. Therefore, Eastvold et al. conducted a meta-analysis of 210 relevant articles (62 met the criteria to be included in the study) examining the effect of an observer on intellectual and academic, attention and
processing speed, executive, learning, memory, and motor task performance. The authors found an overall significant effect size of -0.20 for a physically present TPO and -0.24 for a non-visible (e.g., observing behind a one-way mirror) TPO. Observation through video and audiotaping showed a larger effect size (-0.36); however, inference from this is limited due to the small sample size. Eastvold et al. found that third party observation does have a significant negative effect on neuropsychological task performance, specifically on tasks of attention, learning, and memory. Thus, after review of the literature, it is evident that the presence of an observer is likely to impair significantly certain neuropsychological test performance, regardless of the mode of observation.

**Conation, Third Party Observation, and Task Performance**

Reitan and Wolfson (2000; 2004; 2005) recently brought forth the idea of conation and its relationship to neuropsychological test performance. Conation is referred to as the “ability to marshal and focus intellectual energy that is applied to deal successfully with complex problems that require time to solve” (Reitan & Wolfson, 2000). Conation is thought to involve the use of purposeful, persistent effort and focused intellectual energy to result in optimal performance (Reitan & Wolfson, 2000; 2004). Conation is considered just one component of many that can influence task performance. Others include cognition, motivation, vigilance, and emotion. Reitan and Wolfson (2000; 2004) distinguish between these five important elements. Specifically, conation differs from cognition in that cognition involves the acquisition of information and knowledge while conation includes the effectiveness and persistence of that cognition. Motivation is believed to differ from conation because it is concerned with the role of incentives and needs fulfillment in driving behavior. Vigilance is closely related to conation but
involves a passive process while conation is an active and expressive process. Emotional behavior more obviously differs from conation and is commonly associated with poor performance among competent brain damaged individuals (Reitan & Wolfson, 2000; 2004). Conation is believed to be a crucial connection between intellectual and cognitive capacities and efficient problem solving and task performance (Reitan & Wolfson, 2000).

Reitan and Wolfson (2000; 2004; 2005) performed three studies in which they hypothesized that brain damaged individuals would perform more poorly than controls on tasks that require higher conative functioning and would perform more similarly to controls on tasks low in conative load. Reitan and Wolfson (2000) examined performance on verbal tests (Information, Vocabulary, and Arithmetic from the WAIS), the Speech Sounds Perception Test from the Halstead Reitan Neuropsychological Battery, and the Henmon-Nelson Test of Mental Ability (HNT). It was postulated that the verbal tests would require the least amount of conative ability due to the administration of test items; items are administered one at a time and do not require persistent effort. The Speech Sounds Perception Test was believed to require a moderate amount of conative ability due to the ease of the task but the necessity of focused attention. Lastly, the HNT was hypothesized to require the largest amount of conative ability as it requires 30-minutes of focused attention and effort. Reitan and Wolfson’s (2000) hypotheses were supported in that participants with brain damage showed greater impairment compared to controls as tasks increased in conative demand. Reitan and Wolfson (2004) also provided further evidence that brain damaged individuals perform worse compared to controls as tasks increase in conative demands using (in order of conative ability from lowest to highest) the WAIS Verbal IQ, WAIS Performance IQ, and HNT. Lastly, Reitan and
Wolfson (2005) placed 19 variables from the HRNB in order based on theorized conative ability required for task completion. See Table 1 for a list of the 19 HRNB tests in order from lowest to highest conative ability required. The authors hypothesized that tests requiring a larger amount of conative ability would better differentiate between the brain damaged and control participants. Reitan and Wolson’s (2005) study supported their hypothesis and found that individuals with brain damage performed worse on the tests that required the largest amount of conative ability compared to controls.

Social psychology research established that the presence of an observer often facilitates performance on easy, well-learned tasks but impairs performance on more difficult, complex, or novel tasks (Bond & Titus, 1983). Additionally, literature from the field of neuropsychology have found performance on specific tasks (e.g., memory and learning) to be impaired in the presence of a third party observer, but null results on other tasks (e.g., motor tasks) (Eastvold et al., 2012; Gavett et al., 2005). It is likely that neuropsychological tests fall on a continuum from those low in conative load to those high in conative load. This has been validated with tests from the HRNB along with other intelligence measures (Reitan & Wolfson, 2000; 2004; 2005). Given the different effects (facilitation versus impairment) that observation may have on simple compared to complex tasks, it is possible that these effects also occur for tests high compared to low in conative load. More specifically, the presence of a TPO may cause facilitation on tasks low in conative load and impairment on tasks high in conative load.
Purpose

The purpose of this study is to investigate the effects of third party observation with the use of a one-way mirror on tests that differ in conative load from the HRNB. Two groups of examinees will complete the same tests of varying conative load: one group will have their neuropsychological test performance observed behind a one-way mirror and one group will perform the tests with only the examiner present. In addition, the effects of observation on the California Verbal Learning Test-II (CVLT-II) will also be examined, as this is a popular measure of memory with which the effects of third party observation is unknown. Furthermore, participants will complete both embedded and freestanding measures of symptom validity to ensure that they are putting forth optimal effort throughout testing. The information collected from this study will provide clinicians with additional information about the effects of third party observation on specific assessment measures and assist in decision-making about the presence of a TPO during neuropsychological testing.

Chapter 2

Method

Participants. Following approval from the University at Albany Institutional Review Board, undergraduates were recruited from the university research pool. All participants’ were over age 18 and completed informed consent. Each received one research extra credit upon completion of one hour of participation. A power analysis conducted with the program G*Power (Erdfelder, Faul, & Buchner, 1996) indicated that in order to detect an effect size of -.24 found in the literature (Eastvold et al., 2012) with
80.0% power and alpha set at .05, a total sample size of 60 participants is needed for the first MANCOVA with the five variables that differ in conative load and a total sample size of 64 participants is needed for the second MANCOVA examining performance on six CVLT-II variables.

Data from 100 participants was collected. Participants were questioned about the presence of a learning disability, ADHD, traumatic brain injury (TBI), and/or use of school accommodations. Any participants who self-reported that they experienced any of the above were excluded from the data set. A total of seven cases were excluded from the final analyses. One participant stated that she had ADHD, a learning disability, and history of a TBI; one male reported history of a TBI, one male reported a learning disability and school accommodations; one female stated she had both ADHD and school accommodations in place; one female reported a history of ADHD; one male stated he had a learning disability and school accommodations; and lastly, one female reported history of a TBI.

Participants were also excluded on the basis of their scores on performance validity testing. Examinee performance validity was evaluated using the Medical Symptom Validity Test (MSVT), Reliable Digit Span (RDS), and California Verbal Learning Test-II Forced Choice Recognition (CVLT-II FCR). See Table 2 for the failure criteria for each of these PVTs. This is especially important given the recent findings from An, Zakzanis, and Joorden (2012). These authors found that approximately 56.0% of undergraduate students who participated in their research study failed one or more performance validity test (PVT) in a first session, and that this poor performance was associated with lower scores on neuropsychological assessment measures (An et al.,
Performance invalidity in the current study was identified as failing ≥ two of three PVTs. In this study, no examinee failed two out of three PVTs and thus, no one was identified as putting for invalid performance. In fact, there was only one examinee who failed one PVT. This examinee failed the CVLT-II FCR, but passed all other PVTs. No examinee’s score fell below the RDS or MSVT criteria for invalid performance.

Following exclusion of the seven examinees, the total sample size was 93. There were 58 females (62.4%) and 35 males (37.6%), and 84 were right handed (90.3%). Sixty-four percent were Caucasian, 16.1% African American, 11.80% Asian, and 7.5% Hispanic. Average age was 19.05 (1.16) and years of education were 12.72 (0.97).

Forty-five examinees were assessed with no TPO present (Examiner Only condition) and 48 examinees were observed by a TPO (TPO condition). See Table 3 for detailed demographic information for the total group and each condition (Examiner Only vs. TPO conditions).

**Materials.** Please see Strauss, Sherman, and Spreen, (2006) and Lezak (2012) for a comprehensive discussion of the following assessment instruments.

*Finger Tapping Test* (FTT; Halstead, 1947). The FTT is a measure of manual dexterity and proficiency. This measure can be utilized as a part of a neuropsychological battery to help lateralize and localize an area of injury (Lezak, 2012; Strauss et al., 2006). As can be seen in Table 1, Reitan and Wolfson (2005) ranked the FTT with both the nondominant and dominant hand as tasks from the HRNB that require the lowest conative ability. Thus, the FTT was used in this study as the task with the lowest conative load.
This task is completed with both the dominant and non-dominant hand. Administration time is approximately five minutes. Examinees were instructed to first place their dominant hand on the finger tapper with their palm down and index finger placed on the tapper. The examinee was instructed to tap as quickly as he/she can while moving only his/her index finger and was timed for 10 seconds. Following a practice session, the examinee completed this five consecutive times with his/her dominant hand and then proceeded to do the same with his/her nondominant hand. The slowest and fastest performance on each hand must be within a five-point range; otherwise additional trials (up to 10 total) were administered. Scores on this task included the mean of the five 10-second trials for each hand or the mean of more than five but equal to or less than 10 trials if the mean of the five consecutive trials within a 5-point range was not achieved. Mean raw scores for both the dominant and nondominant hand were used.

*Trail-making Test A and B* (Trails A and Trails B; Reitan, 1955). The Trail-making Test A and B was added to the Halstead’s original tests by Ralph Reitan (Reitan, 1955). This test is purported to measure attention, speed, and mental flexibility as well as sequencing and tracking (Lezak, 2012; Strauss et al., 2006). Trails A involves tracking only one sequence (numbers), while Trails B is more complex and requires tracking and alternating between two sequences (numbers and letters). Reitan and Wolfson (2005) ranked Trails A as requiring moderate conative ability and Trails B as relatively high in conative load (see Table 1). Trails A was examined in this study as a measure of moderate conative ability that lies somewhere between the FTT and the Category Test (CT) and Trails B. Trails B was used as a measure of high conative ability in conjunction with the CT as discussed below. Administration time is approximately five to 10 minutes.
Please see Appendix A for the directions given to examinees for Trails A and Trails B, taken from Strauss et al. (2006). Raw time of completion for both Trails A and Trails B was used.

**Category Test** (*CT*; Halstead, 1947; Hom, 2011; Reitan & Wolfson, 1993). The CT is a test of conceptual reasoning, problem solving, mental flexibility, abstraction ability, and learning (Lezak, 2012; Strauss et al., 2006). There are six unique item sets, with a seventh made up from previous items seen in the first six sets making a total of 208 separate items presented to the examinee. The participant is required to figure out or uncover the principle that runs throughout the entire subtest (Lezak, 2012). The CT was ranked by Reitan and Wolfson (2005) as requiring the highest conative ability second to the examinee’s score on the impairment index. Therefore, this assessment instrument was utilized in this study as the test with the highest conative load.

Silk-Eglit, Gunner, Miele, Lynch, and McCaffrey (2013) found that performance on a new Halstead Category Test computer version (Hom, 2011) does not differ significantly from the standard CT. Therefore, the computerized version of the CT (Hom, 2011) loaded onto a computer running the Windows 7 operating system was administered rather than the standard administration that requires a projector, large wooden, gray box with reflecting mirror and screen, answer panel with four levers, and two carousels of slides. All stimuli were presented on the computer screen. The number of total errors was examined. Administration time is approximately 20 to 30 minutes. See Appendix B for the instructions that were given to examinees during CT administration.
The California Verbal Learning Test-II (CVLT-II; Delis, Kaplan, Kramer, & Ober, 2000; Moore & Donders, 2004). The CVLT-II is a word-list task used to measure verbal learning and memory with semantic grouping strategies (Lezak, 2012; Strauss et al., 2006). TPO effects on this memory measure are yet to be examined, and therefore this study aims to uncover the effects of a TPO on CVLT-II performance. Examinees were administered five learning trials of a 16-item word list, as well as an interference list. They were then asked to complete tasks of immediate and delayed (20-minute) free and cued recall as well as a delayed recognition trial.

Following another 10-minute delay, a forced choice recognition (CVLT-II FCR) trial was utilized as an embedded measure of performance. Moore and Donders (2004) found that a cut-off score of less than 15 correct on the CVLT-II FCR is indicative of performance invalidity (see Table 2). Administration time is approximately 50-minutes, including the 20 and 10-minute delay. Instructions for the CVLT-II are found on the CVLT-II administration forms. Raw number of total correct on all five trials, list B, list A short delay free recall, list A long delay free recall, recognition hits, and recognition false positives were utilized in this study.

Medical Symptom Validity Test (MSVT; Green, 2004). The MSVT was used to assess for invalid performance in addition to the CVLT-II FCR and Reliable Digit Span (RDS; see below). The MSVT is a computerized verbal memory measure. The test is computerized and takes approximately 15 minutes, including the 10-minute delay. Table 2 shows the criteria for failure on the PVT.
Digit Span from the Wechsler Adult Intelligence Scale- 3rd edition (WAIS-III; Wechsler, 1997). Digit Span from the WAIS-III was administered to examinees in order to obtain each examinee’s Reliable Digit Span (RDS; Greiffenstein, Baker, & Gola, 1994). Administration time is approximately five minutes and instructions were taken from the WAIS-III manual. The RDS was utilized as an embedded measure of performance validity. RDS is computed by adding the longest number of digits completed without any errors over two trials both forward and backward (Greiffenstein et al., 1994). Refer to Table 2 for the criteria used to identify performance invalidity.

**Procedure.** Two groups of examinees were administered all of the tests above in the exact same order. Examinees were first administered the MSVT and completed Trails A, Trails B, Digit Span, and the FTT during the 10-minute MSVT delay. After completing the delay on the MSVT, examinees completed the CVLT-II. Examinees were administered the CT during the delay on the CVLT-II. Following completion of the CT, examinees completed the delayed-recall and recognition of the CVLT-II. Examinee’s were then given a demographic information questionnaire and subsequently completed the CVLT-II FCR.

Although the two groups underwent the exact same assessment, half of the examinees were observed by a TPO behind a one-way mirror (i.e., TPO condition), and half performed the neuropsychological tests in front of the examiner (i.e., Examiner Only condition). Those in the TPO condition were introduced to the TPO by the examiner prior to testing so they were aware they were under observation. These examinees were introduced to the TPO behind the one-way mirror so they could see that it was truly a working one-way mirror. They were then brought in to the testing room by the examiner.
The TPO was presented as an expert in the assessment measurements as this has been found to increase social facilitation effects (see Bond & Titus, 1983). In the Examiner Only condition, the one-way mirror was covered with a blind and thus, the examinees were unaware that a one-way mirror was even present in the testing room. Vignettes illustrating the directions given from the examiner to the examinees for both the TPO condition and Examiner Only condition can be found in Appendix C.

**Overview of data collection and analysis.** Data collected included demographic information (age, gender, ethnicity, years of education, and handedness), MSVT scores, Trails A and B time and errors, RDS score, Finger Tapping- Dominant and Nondominant hand mean times, CT subtest and total error scores, and CVLT-II learning trials 1-5, list B, list A short delay free recall, list A long delay free recall, recognition hits, and recognition false positive scores.

First, a between subjects multivariate analysis of covariance (MANCOVA) was used to investigate if there were mean differences in examinee performance on tests differing in conative load (total errors on the CT, time on Trails A, time on Trails B, mean time on FTT- dominant hand, and mean time on FTT-nondominant hand) between the Examiner Only and the TPO conditions while controlling for age, gender, education, handedness, and ethnicity. A second MANCOVA was then utilized to examine performance on a measure of memory, the CVLT-II. Examinee’s total number correct on the five learning trials, list B, list A short delay free recall, list A long delay free recall, number of recognition hits, and number of recognition false positives were examined to see if significant differences existed in performance with the presence of a TPO, while controlling for the same covariates listed above.
Univariate analysis of covariances (ANCOVAs) were also conducted for each dependent variable as follow-up tests to each MANCOVA. These were used to examine if differences in performance exist for each of the separate assessment measures listed above between examinees in the TPO condition and those in the Examiner Only condition. The Bonferroni method for controlling Type 1 error rates for multiple comparisons was utilized for post-hoc tests following each MANCOVA and the alpha level was set at .01 (.05/05) for the conation variables used in the first MANCOVA and .008 (.05/6) for the CVLT-II variables from the second MANCOVA.

Chapter 3

Results

Demographics. Independent sample t-tests indicated that there were no significant differences in age, (t(91) = 0.280, ns), and years of education, (t(91) = 0.302, ns), between the Examiner Only and TPO condition. Chi-square tests were completed to examine differences between the two conditions for gender, handedness, and education. There were no significant differences in gender, (χ²(1, N=93) = 0.782, ns), handedness, (χ²(1, N=93) = 0.205, ns), frequency of Caucasian examinees, (χ²(1, N=93) = 0.200, ns), African American examinees, (χ²(1, N=93) = 2.393, ns), Asian examinees, (χ²(1, N=93) = 0.722, ns), or Hispanic examinees, (χ²(1, N=93) = 0.093, ns), between the two conditions. See Table 3 for descriptive demographic information.

Performance on neuropsychological measures that differ in conative load.

MANCOVA. A between subjects one-way MANCOVA was performed on five dependent variables (DVs): FTT- dominant hand mean time, FTT- nondominant hand
mean time, Trails A time, Trails B time, and the CT total errors. Adjustment was made for five covariates: age, education, gender, handedness, and ethnicity. The independent variable was third party observation (Examiner Only condition vs. TPO condition).

There were no univariate or multivariate outliers. Results of evaluation of assumptions of homogeneity of covariance-variance matrices, linearity, and multicollinearity were satisfactory. Mardia’s (1974) test for multivariate normality indicated that the model violated assumptions of multivariate normality due to significant kurtosis ($\beta_2, p = 43.905, p<.005$). The Shapiro-Wilk’s test was utilized to determine which univariate variables were non-normally distributed. Trails A time, Trails B time, and CT total errors were found to be non-normally distributed. Log transformations were completed on these three variables and the Shapiro-Wilk’s test indicated that all three variables became normally distributed following the transformation. Utilizing the three log-transformed variables along with FTT-dominant and FTT-nondominant (normally distributed without a transformation), Mardia’s (1974) test was no longer significant for violations of multivariate normality. However, when compared to results of the MANCOVA using the non-transformed variables, the MANCOVA with the log-transformed variables presented similar results. Additionally, MANCOVA has been found to be very robust to violations of multivariate normality, especially when the sample size in each group is equivalent and the total sample size is greater than 40 (Ito, 1980; Seo, Kanda, & Fujikoshi, 1995; Tabachnick & Fidell, 2007). Thus, given the total sample size of 93, equal size of examinees per group (48 vs. 45), and the similarity between the log-transformed and non-log transformed findings, MANCOVA with the original, non-log transformed variables remained an appropriate statistical method for
detecting group differences across a variety of DVs. Results from the non-log transformed MANCOVA are discussed below.

With the use of Wilk’s criterion, the combined DV’s were significantly related to gender, $F(5, 82) = 3.319$, $p < .05$, and handedness, $F(5, 82) = 2.372$, $p < .05$, but not to age, $F(5, 82) = 0.610$, ns, ethnicity, $F(5, 82) = 1.143$, ns, education, $F(5, 82) = 0.978$, ns, or the presence of a TPO, $F(5, 82) = 0.892$, ns. There was a modest association between the DVs and gender, partial $\eta^2 = 0.168$, and handedness, partial $\eta^2 = 0.126$. Gender accounted for 16.8% of the variance in the mean differences of the combination of DVs, while handedness accounted for 12.6% of the variance. Small associations were present between the DVs and age, partial $\eta^2 = 0.036$, education, partial $\eta^2 = 0.056$, ethnicity, partial $\eta^2 = 0.065$, and the presence of a TPO, partial $\eta^2 = 0.052$. The presence of a TPO behind a one-way mirror accounted for only 5.2% of the variance in the combination of DV scores. Table 4 presents the means and standard deviations for the DVs included in this first MANCOVA for the two conditions.

**Univariate ANCOVAs.** Univariate ANCOVAs for each DV were conducted as follow-up tests to the MANCOVA. Levene’s test of equality of error variances was not significant for any of the five DVs. Using the Bonferroni method for controlling Type I error rates for multiple comparisons, each ANCOVA was tested at the .01 (.05/5) level when comparing the Examiner Only condition versus the TPO condition because the omnibus F-test was not significant. Due to the significance of the omnibus F-test for both gender and handedness, a Bonferroni correction was not used when examining the univariate ANCOVAs for these covariates.
While controlling for age, gender, education, handedness, and ethnicity, there were no significant differences in scores on the FTT-dominant hand, $F(1, 86) = 0.669$, ns, FTT-nondominant hand, $F(1, 86) = 0.023$, ns, Trails A time, $F(1, 86) = 3.440$, ns, Trails B time, $F(1, 86) = 0.561$, ns, or CT total errors, $F(1, 86) = 0.827$, ns, between the Examiner Only and the TPO condition.

Gender was found to significantly effect mean scores on FTT-dominant hand, $F(1,86) = 6.050, p < .05$, FTT-nondominant hand, $F(1, 86) = 7.823, p < .005$, and Trails B times, $F(1, 86) = 5.647, p < .05$. Gender accounted for 6.6% of the variance in FTT-dominant hand scores (partial $\eta^2 = 0.066$), 8.3% of variance in FTT-nondominant hand scores (partial $\eta^2 = 0.083$), and 6.2% of variance in Trails B time scores (partial $\eta^2 = 0.062$). Specifically, males scored high on the FTT-dominant and FTT-nondominant hand tasks than females and females were quicker to complete Trails B than males. Handedness, however, did not have significant effects on any of the univariate scores. See Table 5 for means and standard deviations for these five DVs organized by gender.

**Performance on the CVLT-II.**

**MANCOVA.** A second between subjects one-way MANCOVA was performed on six DVs from the CVLT-II: total correct on the five learning trials, list B, list A short delay free recall, list A long delay free recall, recognition hits, and recognition false positives. Adjustment was made for the same five covariates as above: age, education, gender, handedness, and ethnicity. The independent variable was third party observation (Examiner Only condition vs. TPO condition).
There were no univariate or multivariate outliers. Results of evaluation of assumptions of linearity and multicollinearity were satisfactory. The assumption of homogeneity of covariance-variance matrices was violated, with Box’s M test being significant (Box’s M= 36.915, p < .05). Additionally, Mardia’s (1974) test for multivariate normality indicated that the model had significant multivariate skew ($\chi^2 = 150.165, p < .001$) and kurtosis ($\beta_2, p = 52.201, p < .05$) and thus, violated assumptions of multivariate normality. The Shapiro-Wilk’s test was used to examine univariate normality. Results indicated that all of the DVs from the CVLT-II were non-normally distributed except for total number correct on the five learning trials. Log transformations were completed on the non-normally distributed variables, however, this transformation did not result in univariate normality for any of these variables. Mardia’s (1974) test was also still significant for multivariate skew ($\chi^2 = 106.531, p < .001$) when using the log-transformed variables. The MANCOVA using these log-transformed variables, however, did meet the assumption of homogeneity of the covariance-variance matrices with Box’s M test being non-significant (Box’s M = 27.403, ns).

Due to the lack of multivariate and univariate normality with the log-transformed variables, square-root transformations were then completed as a second attempt to obtain multivariate normality. The Shapiro-Wilk’s test indicated that square-root transformations were also not able to achieve univariate normality for any of the five DVs. Additionally, Marida’s (1974) test was still significant for multivariate skew ($\chi^2 = 77.037, p < .05$). Use of the square-root transformed variables in the MANCOVA also met the assumption of homogeneity of the covariance-variance matrices with Box’s M being non-significant (Box’s M = 21.969, ns). Due to the inability to achieve
multivariate normality with the use of transformations, the Kruskal-Wallis test was then utilized as a nonparametric method of analysis of variance as it does not assume multivariate normality.

Following inspection of the MANCOVA with the original six DVs, MANCOVA with log-transformed DVs, MANCOVA with square-root transformed DVs, and the non-parametric Kruskal-Wallis test, no differences were found in results between any of the statistical methods. As stated above, MANCOVA has been found to be robust to violations of multivariate normality (Ito, 1980; Seo, Kanda, & Fujikoshi, 1995; Tabachnick & Fidell, 2007). In addition, Box’s M test is notoriously overly sensitivity to violations of the assumption of homogeneity of variance-covariance matrices and is less relevant and useful when sample sizes between groups are equal (Hakstian, Roed. & Lind, 1979; Olson, 1974; Tabachnick & Fidell, 2007). Given the points above as well as the similarity of results between the MANCOVA with the original variables, MANCOVA with the log-transformed variables, MANCOVA with the square-root transformed variables, and Kruskall-Wallis nonparametric test, for ease of interpretation and parsimony, results are discussed from the MANCOVA completed with the original, non-transformed CVLT-II variables.

With the use of Wilk’s criterion, the combined DV’s were not significantly related to the presence of a TPO, $F(6, 81) = 1.620$, ns, age, $F(6, 81) = 0.173$, ns, education, $F(6, 81) = 1.232$, ns, handedness $F(6,81) = 0.255$, ns, or ethnicity, $F(6, 81) = 0.691$, ns. The combined DV’s, however, were significantly related to gender, $F(6, 81) = 2.398$, $p < .05$. There was a modest association between the DVs and gender, partial $\eta^2 = 0.151$, as well as the DVs and the presence of a TPO, partial $\eta^2 = 0.107$. Gender
accounted for 15.1% of the variance in the mean differences of the combination of DVs and the presence of a TPO accounted for 10.7% of the variance in the combination of CVLT-II scores. Small associations were present between the DVs and age, partial $\eta^2 = 0.039$, education, partial $\eta^2 = 0.084$, handedness, partial $\eta^2 = 0.019$, and ethnicity, partial $\eta^2 = 0.041$. Table 6 presents the means and standard deviations for the DVs included in this second MANCOVA for each condition.

**Univariate ANCOVAs.** Univariate ANCOVAs for each of the six DVs were conducted as follow-up tests to the second MANCOVA. Levene’s test of equality of error variances was only significant for one variable: recognition false positives; however, as stated above, ANCOVA is relatively robust to univariate normality (Ito, 1980). Since the omnibus F-test was not significant for the effect of a TPO, Bonferroni method for controlling Type I error rates for multiple comparisons was used and each ANCOVA was tested at the .008 (.05/6) level when comparing the Examiner Only versus the TPO conditions. While controlling for age, gender, education, handedness, and ethnicity, there were no significant differences in scores on total correct for the five trials, $F(1,86) = 2.708$, ns, list b, $F(1, 86) = 0.701$, ns, list A short delay free recall, $F(1, 86) = 1.528$, ns, list A long delay free recall, $F(1, 86) = 0.955$, ns, recognition hits, $F(1, 86) = 0.006$, ns, and recognition false positives, $F(1, 86) = 4.715$, ns, between the Examiner Only and TPO conditions.

The omnibus effect of gender, however, was significant and therefore, a Bonferroni correction for multiple comparisons was not used when examining the univariate effects of this covariate. Gender was found to have a significant effect on scores from list B, $F(1, 86) = 8.406$, $p < .01$, with females ($m = 6.41$, $sd = 1.68$) scoring
higher than males (m = 5.54, sd = 1.63). Gender accounted for 8.9% of the variance in scores on list B, partial $\eta^2 = 0.089$. Gender was not found to have a significant effect on any of the other CVLT-II scores. Education, age, handedness, and ethnicity were not found to have any significant effect on any of the six CVLT-II DVs.

Chapter 4

Discussion

The purpose of this study was two-fold. First, this study sought to examine the effects of third party observation with the use of a one-way mirror on neuropsychological assessment measures from the HRNB that vary in conative load. Five tests from the HRNB were utilized as a sample of tests with varying conative load: the FTT-nondominant hand, the FTT-dominant hand, Trails A, Trails B, and the CT. Table 1 displays the rank of these tests in order from lowest to highest conative load (Reitan & Wolfson, 2005). FTT-nondominant hand and FTT-dominant hand were used in this study to represent tests that require low conative ability. Trails A was utilized as an assessment measure that required moderate conative ability, and Trails B and the CT were used as they are tests ranking higher in conative load.

Historical literature from the social psychology field has determined that the presence of an observer results in a social facilitation effect on easy or well-learned tasks and a social loafing effect on more complex or novel tasks (Bond & Titus, 1983). Neuropsychological literature has also found that performance is improved on more difficult measures of memory and learning, but unaffected on other tasks, such a motor tasks, in the presence of an observer (Eastvold et al., 2012; Gavett et al., 2005). Thus, it
was hypothesized that performance on neuropsychological tests requiring higher conative ability (i.e., Trails B and the CT) would be impaired in the presence of a TPO behind a one-way mirror, while performance on tests requiring lower conative ability (FTT-nondominant hand and FTT-dominant hand) would be facilitated or unaffected.

Secondly, prior research from both the social psychology (Deffenbacher, Platt, & Williams, 1974; Geen, 1971; Geen, 1973) and neuropsychology fields (Constantinou et al. 2002; Constantinou et al., 2005; Eastvold et al., 2012; Gavett et al., 2005; Keher et al. 2012; Lynch, 2005) has found that performance on various measures of learning and memory are impaired in the presence of a second observer or TPO regardless of mode of observation (e.g., TPO present, video recording, or audio recording). Therefore, an important component of this study was to examine the effects of observation behind a one-way mirror on a measure of memory, the CVLT-II, as the effects of a TPO on this task were yet to be determined.

Undergraduate students were recruited through the University Research Pool and administered the FTT- nondominant and dominant hand, Trail-making Test A and B, CT, and CVLT-II under two conditions. In the first condition, examinees were administered the assessment instruments in the presence of only the examiner. In the second condition, examinees were administered the neuropsychological tests in the presence of both the examiner and a TPO seated behind a one-way mirror. Examinees were introduced to the TPO and shown the one-way mirror prior to the assessment and were aware their performance was being observed by an individual who had expert knowledge of the tests being administered. Three tasks: the MSVT, Digit Span RDS, and CVLT-II FCR were used as measures of performance validity. No examinees were found to be putting forth
invalid effort as determined by failing two or more of these measures. Additionally, examinees in both group were screened for learning disability, ADHD, TBI, and the utilization of school accommodations. Seven examinees endorsed that they suffered from a learning disability, ADHD, TBI and/or utilized school accommodations and were subsequently removed from all further analyses resulting in a full sample of N = 93 (Examiner Only condition = 45 and TPO condition = 48).

**Key findings.** A MANCOVA was conducted to determine if third party observation behind a one-way mirror affected examinee performance on tests of varying conative load while controlling for age, education, gender, handedness, and ethnicity. Results indicated that, dissimilarly from previous research (see Eastvold, 2012; Gavett et al., 2005), third party observation had no statistically significant effect on any of the assessment measures (FTT, Trail-making Test, and CT). Specifically, it was thought that examinees being observed by a third party would perform better on FTT-dominant hand, FTT-nondominant hand, and possibly Trails A and worse on Trails B and the CT given their relative conative load (Reitan & Wolfson, 2005). Additionally, results from prior social psychology research indicated facilitation on easy tasks and impairment on more difficult tasks when being observed (see Bond & Titus, 1983). However, results from this study suggested that observation behind a one-way mirror did not have a statistically significant influence on examinee performance, regardless of the conative load of the neuropsychological measure as measured by statistical significance. Observation by a third party behind a one-way mirror accounted for only 5.2% of the variance in the combination of mean differences on FTT-nondominant hand time, FTT-dominant hand time, Trails A time, Trails B time, and CT total errors.
In fact, statistically significant mean differences on these neuropsychological assessment measures were found only as a result of gender and handedness. Gender accounted for 16.8% of the variance in the combination of mean difference scores on the five conation variables. Univariate analyses revealed that males performed better than females on both the FTT-nondominant and FTT-dominant task and gender accounted for 8.3% and 6.6% of the variance in these scores, respectively. This finding is not surprising, as prior research has found that, in general, men perform better than women on the FTT-nondominant and FTT-dominant tasks. The disparity in performance between genders is likely due to phenotypic differences and larger hand size rather than differences in neuropsychological functioning (Dodrill, 1979; Ruff & Parker, 1993). Alternatively, females showed better performance and were quicker to complete Trails B than males. Gender accounted for 6.2% of the variance in scores on Trails B. This is surprising given that prior research has not found a difference between male and female performance on Trails B in normative samples (Giovagnoli, Del Pesce, Mascheroni, Simonvelli, Laiacona, & Capitani, 1996; Tombaugh, 2004). Univariate analyses revealed no differences in performance between examinees for the TPO vs. Examiner only conditions or other covariates (handedness, ethnicity, age, or education) on any of the measures.

A second MANCOVA was utilized to determine if the presence of a TPO behind a one-way mirror affected performance on a frequently used neuropsychological measure of memory, the CVLT-II, while controlling for age, education, gender, handedness, and ethnicity. Six scales from the CVLT-II were used as DVs in the MANCVOA as measures of learning, free recall, and recognition memory: total correct on trials one
through five, list B, list A short delay free recall, list A long delay free recall, recognition hits, and recognition false positives. Similar to above, contrary to what was expected, the omnibus effect of observation by a third party behind a one-way mirror was not statistically significant. The presence of a TPO behind a one-way mirror did not have a statistically significant effect on the combination of mean difference scores from the CVLT-II between the two conditions. The presence of a TPO accounted for more variance in the CVLT-II scores than the conation variables, accounting for 10.7% of the total variance, however, this effect did not reach statistical significance.

Once again, statistically significant mean differences between the CVLT-II scores were found only for the effect of gender. Gender accounted for 15.1% of the variance in the six CVLT-II learning and memory scores. No omnibus effect was statistically significant for any of the other covariates. Univariate analyses revealed that females had higher scores than males on list B of the CVLT-II and gender accounted for 8.9% of the variance in these scores. This finding is similar to prior research that has provided evidence that females tend to perform better on list B than males (Normal, Evans, Miller, & Heaton, 2000; Wiens, Tindall, & Crossen, 1994). Interestingly, though, Norman et al. (2000) and Wiens et al. (1994) found that females performed better than males on multiple aspects of the CVLT, not only list B, including total words recalled on trials one through five, trail 1, trial 5, short delay free recall, and long delay free recall, while the current study only established an effect of gender on list B. It is important to note, however, that these studies examined performance on the CVLT, not CVLT-II. Nevertheless, given the similarity between the two versions it would be expected that results from research and validation studies would be highly analogous (Delis et al.,
No additional univariate effects for gender, TPO condition, or any of the other covariates (handedness, age, education, or ethnicity) were found to be statistically significant.

An interesting secondary finding involved the use of PVTs in this sample of undergraduate volunteers in order to detect performance invalidity. The current study found that no examinees were identified as putting forth invalid performance when invalid performance was defined as failing two or more of three PVTs (CVLT-II FCR, Digit Span RDS, and MSVT). This finding was highly dissimilar from a prior study that discovered approximately 56% of undergraduate research volunteers failed at least one PVT (An et al., 2012). Dissimilarly from the current study, performance invalidity was identified as failing only one PVT in the An et al. (2012) study. However, even if the An et al. (2012) criteria was utilized to define performance invalidity in the current study, only one undergraduate would have been flagged as putting forth invalid performance. This examinee failed the CVLT-II FCR with a score of 14 out of 16, but passed all other PVTs and thus was identified as performance valid for data analyses in this study.

Differences in PVTs administered between the two studies may have accounted for these differences. An et al. (2012) assessed for invalid performance using the Test of Memory Malingering (TOMM; Tombaugh, 1996), Victoria Symptom Validity Test (VSVT; Slick, Hopp, & Strauss, 1997), and Dot Counting Test (DCT; Boone, Lu, & Herzberg, 2002). No participants in An et al.’s (2012) study failed performance validity criteria on the TOMM, but 65% and 45% failed the VSVT and DCT, respectively. It is possible that the PVTs, specifically the DCT and VSVT, used in An et al.’s (2012) study are more sensitive indicators of performance validity rather than those used in the current study.
study. For example, Strauss, Slick, Levy-Bencheton, Hunter, MacDonald, and Hultsch (2002) found that when comparing the effectiveness of the VSVT, RDS, and computerized DCT, the VSVT was best at detecting invalid performance, followed by the RDS then computerized DCT. Alternatively, the CVLT-FRC cut-off of less than 15 out of 16 correct has been found to have 89% agreement with the TOMM (Moore & Donders, 2004). The TOMM in An et al.’s (2012) study did not identify any individuals as putting forth invalid performance. Given the similarity in classification between the CVLT-II FCR and TOMM, it is possible that the CVLT-II FCR did not accurately identify examinees with invalid performance in the current study (Moore & Donders, 2004). Furthermore, the MSVT has been found to show very high sensitivity and specificity in the identification of invalid performance, although it may not be as sensitive as longer free-standing PVTs such as the Word Memory Test (Carone, 2009). Consequently, it is possible that administration of a longer or more sensitive free-standing PVT may have resulted in a greater identification of invalid performance. It is unlikely, however, that the choice of administered PVTs completely explains the large difference in frequency of invalid performance between the current study and An et al. (2012) study (i.e., a 56% difference).

Another factor that may have affected the difference in the undergraduates’ performance between the two studies is their motivation level. In the current study, 65 out of the 100 participants were assessed in the first half of the semester. These undergraduates may have been more motivated and proactive to receive credit and participate in the research study than undergraduates attending sessions closer to the end of the semester. In fact, 22 undergraduates signed up but then did not show for their
session throughout the semester. Sixteen of the 22 students were those who signed up in the second half of the semester (following spring break) possibly providing some evidence for the idea that students who sign up earlier in the semester are more motivated to complete the research studies. Furthermore, multiple students voiced their interest in the research study and the goals of the study both prior to and following testing. The majority of the students were required to participate in multiple research studies as a component of the introductory psychology class. These students often told the examiner that the current study was the most interesting of the studies of which they have participated. This increased level of interest may have led to a decreased frequency of performance invalidity in the sample.

**Clinical implications.** Observation by a third party during a neuropsychological evaluation can occur in either a clinical or forensic context and can include parents of the examinee, trainees, supervisors, other psychologists or health care providers, attorneys, and/or legal assistants (McSweeny et al., 1998). Observation can be completed directly or indirectly and involve the use of audio and video recorders, or a one-way mirror. The presence of a TPO during a neuropsychological evaluation has been called into question for various reasons, including decreased validity and reliability of the assessment (Axelrod et al., 2000), increased distraction (Antastasi, 1988; Lynch & McCaffrey, 2004), decreased value of standardized normative data (Howe & McCaffrey, 2012; Lynch & McCaffrey, 2004; McCaffrey et al., 1996), ethical violations due to the possibility of violations of test security and test misuse (AACN, 2001; Duff & Fisher, 2005; McCaffrey et al., 1996), and lastly increased likelihood of breaches of confidentiality (Shealy et al., 2008). Due to these potential issues, three organizations, AACN (2001), NAN (2000),
and the American Psychological Association’s Committee on Psychological Tests and Assessment (2007), have put forth consensus statements regarding the presence of a TPO during an assessment. These organizations stress the importance of limiting the presence of TPOs during evaluations, but recognize that in some instances (e.g., training), the presence of a TPO is unavoidable. Accordingly, in these unavoidable situations, the impact of a TPO should be minimized to the greatest degree possible. The American Psychological Association’s Committee on Psychological Tests and Assessment (2007) recommends the use of a one-way mirror for observation as a method that results in the least interference, although AACN (2001) states that indirect means of observation still affect the examinee and are unacceptable alternatives.

In the current study, no statistically significant effects as a result of third party observation behind a one-way mirror were found on any of the tests utilized, regardless of whether the measures were difficult, easy, high in conative load, or low in conative load. Examinees’ performance on FTT-nondominant hand, FTT-dominant hand, Trails A, Trails B, the CT, and the CVLT-II did not differ as a result of condition. Examinees introduced to a TPO behind a one-way mirror did not show improved performance on the tasks that required less conative ability and decrements in performance on tasks that required more conative ability or the CVLT-II, like expected.

These findings are surprising given that they are in contrast to prior research in the field of neuropsychology that have found impairment in performance on neuropsychological tasks for individual’s who are being observed by a third party either directly or indirectly (see Eastvold et al. 2012; Gavett et al., 2005). In one meta-analyses, performance behind a one-way mirror even had a medium size effect of -.24 (Eastvold et
al., 2012). Replication is necessary before any conclusions can be reached or recommendations followed based on the results of the current study given the findings from prior social psychology and neuropsychology research. Further research is necessary to determine if observation behind a one-way mirror truly has null effects on examinee’s performance on these specific neuropsychological tasks or if other variables may be present in this study that are suppressing the effects of observation behind a one-way mirror. Thus, due to the breadth of prior research evidencing the consequences that third party observation may have on neuropsychological test performance in addition to the various issues associated with observation discussed above, TPOs should still be avoided to the maximum extent possible.

**Limitations and conclusions.** The population from which examinees in this study were sampled was undergraduate students receiving class or extra credit for participation in research studies. The findings from this study may not generalize to neuropsychological evaluations in which a third party is present with clinical and/or forensic populations. It is possible that the consequences associated with a clinical and/or forensic evaluation are greater and subsequently result in higher levels of examinee anxiety or activation in the presence of a third party behind a one-way mirror, leading to larger effects on neuropsychological task performance. Additionally, individuals who have had a brain injury or neurological damage may have more difficulty on neuropsychological tasks and have increased feelings of judgment in the presence of an observer compared to undergraduate students. However, prior research has utilized undergraduate samples and evidenced consequences of observation on neuropsychological performance (see Eastvold et al., 2012; Gavett et al., 2005).
Nonetheless, the study should be cross-validated with clinical groups and varying neuropsychological tests.

Another variable that may have resulted in null effects is the characteristics of the TPO and examiner. Although the TPO was presented as an expert and professional in her mannerisms, dress, and introductions, she was still a young female and may not have exuded the effect of an older, more established professional. Additionally, a particularly likeable TPO and examiner likely does not have similar effects on examinee performance as certain professionals who may observe an evaluation, such as an oppositional attorney. Although previous research has found that increased anxiety due to observation is not necessary to detect effects of social facilitation or loafing, it may be interesting to include a measure of anxiety to compare levels of anxiety in differing conditions (e.g., no observer, direct observer, audio/video-recording, and one-way mirror) with various types of observers (e.g., male vs. female, older vs. younger, and friendly vs. stoic; Geen & Gange, 1977; Zajonc, 1965). This would be useful in order to determine which conditions lead to increases in anxiety among examinees and whether this increase in anxiety has any influence on task performance.

The purpose of this study was to determine the effects of third party observation with the use of a one-way mirror on tasks that varied in conative load as well as a memory measure that has yet to be examined. There are multiple ethical, professional, and testing validity concerns associated with the use of a TPO during a neuropsychological evaluation. Prior research has found performance decrements on more difficult tasks of memory, learning, executive functioning, and other cognitive domains. This research revealed that the use of a one-way mirror did not have
statistically significant effects on neuropsychological task performance, regardless of the conative load of the assessment tool. In addition, performance on the CVLT-II, a learning and memory measure, was also statistically similar between examinees observed behind a one-way mirror and examinees assessed with only the examiner present. These findings, however, are inconsistent with prior research evidencing the effect of third party observation on task performance. As there is a multitude of evidence exhibiting the effects of TPO on neuropsychological measures, it is still best to avoid the presence of a TPO in any clinical and/or forensic evaluation. Future research is needed to confirm or refute the findings from this study demonstrating no statistically significant effects on these specific neuropsychological tasks when observation is completed behind a one-way mirror.
References


Norman, M. A., Evans, J. D., Miller, S. W., & Heaton, R. K. (2000). Demographically


<table>
<thead>
<tr>
<th>Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finger Tapping- Nondominant Hand</td>
</tr>
<tr>
<td>Finger Tapping- Dominant Hand</td>
</tr>
<tr>
<td>Bilateral Auditory Stimulation</td>
</tr>
<tr>
<td>Bilateral Visual Stimulation</td>
</tr>
<tr>
<td>Verbal IQ</td>
</tr>
<tr>
<td>Bilateral Tactile Stimulation</td>
</tr>
<tr>
<td>Tactile Finger Recognition</td>
</tr>
<tr>
<td>Tactile Form Recognition</td>
</tr>
<tr>
<td>Rhythm Test</td>
</tr>
<tr>
<td>Speech-sounds Perception Test</td>
</tr>
<tr>
<td>Trail-making Test- Part A</td>
</tr>
<tr>
<td>Performance IQ</td>
</tr>
<tr>
<td>Tactual Performance Test-Memory</td>
</tr>
<tr>
<td>Tactual Performance Test-Localization</td>
</tr>
<tr>
<td>Finger-tip Number Writing</td>
</tr>
<tr>
<td>Trail-Making Test- Part B</td>
</tr>
<tr>
<td>Tactual Performance Test- Time</td>
</tr>
<tr>
<td>Category Test</td>
</tr>
<tr>
<td>Impairment Index</td>
</tr>
</tbody>
</table>

*As found in Reitan and Wolfson (2005)
Table 2. *Criteria for Failure on Measures of Performance Validity.*

<table>
<thead>
<tr>
<th>PVT</th>
<th>Failure Determined by:</th>
</tr>
</thead>
<tbody>
<tr>
<td>RDS</td>
<td>Score of &lt; 7 correct&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>CVLT-II FCR</td>
<td>Score of &lt; 15 correct&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>MSVT</td>
<td>&lt; 85% correct on Immediate Recognition,</td>
</tr>
<tr>
<td></td>
<td>Delayed Recognition or Consistency Index&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> (Babikian, Boone, Lu, & Arnold, 2006; Schroeder, Twumasi-Ankrah, Baade, & Marshall, 2012)
<sup>b</sup> (Moore & Donders, 2004)
<sup>c</sup> (Green, 2004)
Table 3. Demographic Information for Total Group, Examiner Only Condition, and TPO Condition.

<table>
<thead>
<tr>
<th></th>
<th>Total Group n=93</th>
<th>Examiner Only Condition n=45</th>
<th>TPO Condition n=48</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m \ (sd)$</td>
<td>19.05 (1.16)</td>
<td>19.09 (1.28)</td>
<td>19.02 (1.06)</td>
</tr>
<tr>
<td>Range</td>
<td>18-23</td>
<td>18-23</td>
<td>18-22</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m \ (sd)$</td>
<td>12.72 (0.97)</td>
<td>12.69 (1.02)</td>
<td>12.75 (0.93)</td>
</tr>
<tr>
<td>Range</td>
<td>12-15</td>
<td>12-15</td>
<td>12-15</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M/F)</td>
<td>35/58</td>
<td>19/26</td>
<td>16/32</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Caucasian</td>
<td>64.5</td>
<td>62.2</td>
<td>66.7</td>
</tr>
<tr>
<td>% African American</td>
<td>16.1</td>
<td>22.2</td>
<td>10.4</td>
</tr>
<tr>
<td>% Asian</td>
<td>11.8</td>
<td>8.9</td>
<td>14.6</td>
</tr>
<tr>
<td>% Hispanic</td>
<td>7.5</td>
<td>6.7</td>
<td>8.3</td>
</tr>
<tr>
<td><strong>Handedness</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R/L)</td>
<td>84/9</td>
<td>40/5</td>
<td>44/4</td>
</tr>
</tbody>
</table>
Table 4. Means and Standard Deviations for Five DVs Differing in Conative Load Included in the First MANCOVA for the Examiner Only Condition and TPO Condition.

<table>
<thead>
<tr>
<th></th>
<th>Examiner Only Condition</th>
<th>TPO Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTT- dominant hand</td>
<td>46.63 (6.90)</td>
<td>47.67 (6.03)</td>
</tr>
<tr>
<td>m (sd)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTT- nondominant hand</td>
<td>44.34 (5.28)</td>
<td>44.29 (5.92)</td>
</tr>
<tr>
<td>m (sd)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trails A time</td>
<td>22.31 (5.76)</td>
<td>20.27 (5.04)</td>
</tr>
<tr>
<td>m (sd)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trails B time</td>
<td>52.02 (17.45)</td>
<td>49.29 (11.24)</td>
</tr>
<tr>
<td>m (sd)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT total errors</td>
<td>32.80 (15.73)</td>
<td>30.08 (15.91)</td>
</tr>
<tr>
<td>m (sd)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5. *Means and Standard Deviations on the Five DVs Differing in Conative Load Organized by Gender.*

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTT- dominant hand</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m \ (sd)$</td>
<td>49.09 (5.92)</td>
<td>46.00 (6.53)</td>
</tr>
<tr>
<td>FTT- nondominant hand</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m \ (sd)$</td>
<td>46.43 (5.88)</td>
<td>42.98 (5.03)</td>
</tr>
<tr>
<td>Trails A time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m \ (sd)$</td>
<td>22.20 (7.11)</td>
<td>20.69 (4.16)</td>
</tr>
<tr>
<td>Trails B time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m \ (sd)$</td>
<td>55.11 (19.14)</td>
<td>47.90 (10.19)</td>
</tr>
<tr>
<td>CT total errors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m \ (sd)$</td>
<td>29.31 (16.06)</td>
<td>32.66 (15.64)</td>
</tr>
</tbody>
</table>
Table 6. Means and Standard Deviations for Six DVs from the CVLT-II Included in the Second MANCOVA for the Examiner Only Condition and TPO Condition.

<table>
<thead>
<tr>
<th></th>
<th>Examiner Only Condition</th>
<th>TPO Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Correct on Five Trials</td>
<td>56.93 (9.63)</td>
<td>54.19 (7.96)</td>
</tr>
<tr>
<td>List B</td>
<td>5.91 (1.79)</td>
<td>6.25 (1.62)</td>
</tr>
<tr>
<td>List A Short Delay Free Recall</td>
<td>12.27 (2.59)</td>
<td>11.65 (2.55)</td>
</tr>
<tr>
<td>List A Long Delay Free Recall</td>
<td>12.71 (2.29)</td>
<td>12.25 (2.28)</td>
</tr>
<tr>
<td>Recognition Hits</td>
<td>15.11 (1.27)</td>
<td>15.08 (1.15)</td>
</tr>
<tr>
<td>Recognition False Positives</td>
<td>0.67 (0.91)</td>
<td>1.31 (1.84)</td>
</tr>
</tbody>
</table>
Appendix A

Instructions for the Trail-making Test (taken from Strauss et al., 2006)

Part A

Sample A. When ready to begin the test, place the Part A test sheet in front of the subject, give the subject a pencil, and say: On this page (point) are some numbers. Begin at number 1 (point to “1”) and draw a line from 1 to 2, (point to “2”), 2 to 3 (point to “3”), 3 to 4 (point to “4”), and so on, in order, until you reach the end (pointing to the circle marked “END”). Draw the lines as fast as you can. Do not lift the pencil from the paper. Ready! Begin!

If the subject makes a mistake on Sample A, point it out and explain it. The following explanations of mistakes are acceptable:

1. You started with the wrong circle. This is where you start (point to “1”).

2. You skipped this circle (point to the one omitted). You should go from number 1 (point) to 2 (point), 2 to 3 (point), and so on, until you reach the circle marked “END” (point).

3. Please keep the pencil on the paper, and continue right on to the next circle.

After the mistake has been explained, the examiner marks out the wrong part and says: Go on from here (point to the last circle completed correctly in the sequence).

If the subject still cannot complete Sample A, take the subject’s hand and guide the pencil (eraser end down) through the trail. Then say: Now you try it. Put your pencil,
point down. Remember, begin at number 1 (point), and draw a line from 1 to 2 (point to “2”), 2 to 3 (point to “3”), 3 to 4 (point to “4”), and so on, in order until you reach the circle marked “END” (point). Do not skip around but go from one number to the next in the proper order. If you make a mistake, mark it out. Remember, work as fast as you can. Ready! Begin!

If the subject succeeds this time, go on to Part A of the test. If not, repeat the procedure until the subject does succeed or it becomes evident that he or she cannot do it.

If the subject completes the sample item correctly, and in a manner which shows that he or she knows what to do, say: Good! Let’s try the next one. Turn the page and give Part A of the test.

Say, On this page are numbers from 1 to 25. Do this the same way. Begin at number 1 (point) and draw a line from 1 to 2 (point to “2”), 2 to 3 (point to “3”), 3 to 4 (point to “4”), and so on, in order until you reach the end (point). Remember, work as fast as you can. Ready! Begin!

Start timing. If the subject makes an error, call it to his or her attention immediately, and have the subject proceed from the point where the mistake occurred. Do not stop timing.

If the examinee completes Part A without error, remove the test sheet. Record the time in seconds. Errors count only in the increased time of performance. Then say: “That’s fine. Now we’ll try another one.” Proceed immediately to Part B, sample.
Part B

Sample B. Place the test sheet for Part B, sample side up, flat on the table in front of the examinee, in the same position as the sheet for Part A was placed. Point with the right hand to the sample and say: On this page are some numbers and letters. Begin at number 1 (point) and draw a line from 1 to A (point to “A”), A to 2 (point to “2”), 2 to B (point to “B”), B to 3 (point to “3”), 3 to C (point to “C”), and so on, in order until you reach the end (point to circle marked “END”). Remember, first you have a number (point to “1”), then a letter (point to “A”), then a number (point to “2”), then a letter (point to “B”), and so on. Draw the lines as fast as you can. Ready! Begin!

If the subject makes a mistake on Sample B, point it out and explain it. The following explanations of mistakes are acceptable:

1. You started with the wrong circle. This is where you start (point to “1”).

2. You skipped this circle (point to the one omitted). You should go from 1 (point) to A (point), A to 2 (point), 2 to B (point), B to 3 (point), and so on until you reach the circle marked “END” (point). If it is clear that the subject intended to touch the circle but missed it, do not count it as an omission, but caution him or her to touch the circle.

3. You only went as far as this circle (point). You should have gone to the circle marked “END” (point).

4. “Please keep the pencil on the paper and go right on to the next circle.”
After the mistake has been explained, the examiner marks out the wrong part and says: Go on from here (point to the last circle completed correctly in the sequence).

If the subject still cannot complete Sample B, take the subject’s hand and guide the pencil (eraser end down) through the circles. Then say: Now you try it. Remember, you begin at number 1 (point) and draw a line from 1 to A (point to “A”), A to 2 (point to “2”), 2 to B (point to “B”), B to 3 (point to “3”), and so on until you reach the circle marked “END” (point). Ready! Begin!

If the subject succeeds this time, go on to Part B of the test. If not, repeat the procedure until the subject does succeed or it becomes evident that he or she cannot do it.

If the subject completes the sample item correctly, say: Good. Let’s try the next one. Turn the page over and proceed immediately to Part B, and say: On this page are both numbers and letters. Do this the same way. Begin at number 1 (point) and draw a line from 1 to A (point to “A”), A to 2 (point to “2”), 2 to B (point to “B”), B to 3 (point to “3”), 3 to C (point to “C”), and so on, in order, until you reach the end (point to circle marked “END”). Remember, first you have a number (point to “1”), then a letter (point to “B”), and so on. Do not skip around, but go from one circle to the next in the proper order. Draw the lines as fast as you can. Ready! Begin!

Start timing. If the subject makes an error, immediately call it to his or her attention and have the subject proceed from the point at which the mistake occurred. Do not stop timing.

If the subject completes Part B without error, remove the test sheet. Record the time in seconds. Errors count only in the increased time of performance.
Appendix B

Category Test Instructions (Strauss et al., 2006)

Say to the patient: On this screen you are going to see different geometrical figures and designs. Something about the pattern on the screen will remind you of a number between 1 and 4. On the keyboard in front of you (pointing) the keys are numbered 1, 2, 3, and 4. You are to press down on the key that is the same number that the pattern on the screen reminds you of. That is, if the picture on the screen reminds you of the number 1, pull key number 1. If the picture on the screen reminds you of the number 2, pull key number 2. And so on. For example, what number does this remind you of?

Put on the first slide. If the subject says “one,” ask the subject which key he or she should press. After the subject has pressed the number 1 key, say: The bell you have just heard tells you that you got the right answer. Every time you have the right answer you will hear the bell ring. Instruct the subject to try one of the other keys to find out what happens when an incorrect key is pressed. Then say: The buzzer is what you hear when you have the wrong answer. In this way, you will know each time whether you have the right or wrong answer. However, for each picture on the screen you get only one choice. If you make a mistake we just go right on to the next picture.

Proceed with Subtest I. Say: Now which key would you pick for this picture? After Subtest I, say: That was the end of the first subtest. This test is divided into seven subtests. In each subtest there is one idea or principle that runs throughout the entire subtest. Once you have figured out what the idea or principle in the subtest is, by using this idea you will get the right answer each time. Now we are going to begin the second subtest.
The idea in it may be the same as in the practice set, or it may be different. We want you to figure it out.

Proceed with Subtest II. After Subtest II, say: That was the end of the second subtest and, as you probably noticed, you don’t necessarily have to see a number to have a number suggested to you. You saw squares, circles, and other figures. Also, as you may or may not have noticed, in each of these subtests, there is one idea or principle that runs throughout. Once you figure out the idea, you continue to apply it to get the right answer.

Now we are going to start the third subtest. The idea may be the same as the last one or it may be different. I want to see if you can figure out what the idea is and then use it to get the right answer. Remember, the idea remains the same throughout the sub-test. I will tell you when we complete one subtest and are ready to begin a new one.

Proceed with Subtest III. In Subtest IV, after slide #6 (the first slide without numbers), say: This is still the same group, but now the numbers are missing. The principle is still the same. After Subtests III, IV, and V, say: That was the end of that subtest. Now we are going to begin the next one. The idea in it may be the same as the last one or it may be different. We want you to figure it out.

After Subtest VI, say: In the last subtest there is no one idea or principle that runs throughout the group because it is made up of items you have already seen in preceding subtests. Try to remember what the right answer was the last time you saw the pattern and give that same answer again.
Appendix C

Examiner Only Condition

“Today you are going to be completing a variety of tests that will assess various aspects of your cognitive functioning such as learning, memory, attention, motor speed, and executive functioning. Some of these tests will be easy for you and some will be difficult. You are not expected to answer all of the questions correctly, only that you put forth your best effort. In total, this study will take approximately one hour of your time. You will be receiving one Sona-Systems research credit. I will provide you with directions regarding the procedure at the beginning of each test, as well as prompts throughout most of the assessments, but please let me know at any time if you have any further questions. Do you have any questions? Let’s get started.”

TPO Condition

During the TPO condition, the examinee will be introduced to the TPO prior to testing. The introduction will proceed as follows:

“This is (insert name). She has been trained extensively in the cognitive measures that we will be using today and will be observing the assessment behind this one-way mirror for the hour that you are here”.

The TPO will then say hello, shake hands with the participant and then retreat behind the one-way mirror. After the TPO has settled behind the one-way mirror the directions to the examinee will be the same as those written above in the Examiner Only condition.