UNIQUE VARIABILITY BETWEEN THE
DEAN-WOODCOCK EMOTIONAL STATUS
EXAMINATION AND THE MMPI

A DISSERTATION
SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE
DOCTOR OF PHILOSOPHY

BY
MEGAN M. MORSE

DISSERTATION ADVISOR: DR. RAYMOND DEAN

BALL STATE UNIVERSITY
MUNCIE, INDIANA

MAY 2013
DEDICATION

This work is dedicated to my love. You provided support in every way and gave me the courage and motivation to take the steps necessary to complete this process. I cannot imagine my life without you, and my deepest desire is to be yours forever. So will you be mine… forever?
ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to Dr. Raymond Dean who served as the chair of my dissertation committee and who provided me with invaluable guidance, advice, and support. I would also like to thank Dr. Eric Pierson, Dr. Sharon Paulson, and Dr. David LeBlanc, my doctoral committee, for their input and their patience. I am extremely grateful to Dr. Holmes Finch for his assistance with the interpretation of the statistical analysis of this work. I am forever in debt to Linda Jones and Penny Nowaczyk for their help in getting through this process.

I would like to take this opportunity to acknowledge the endless support and direction provided by Dr. Kenneth McCoy. I will never take for granted that the doctor I become will reflect his dedication to student training and supervision.

I wish to extend my most sincere love and gratitude to my family and friends for their unwavering support. They have been invaluable throughout this process and have believed in me even when I have not believed in myself. Each of them has a special place in my heart.
TABLE OF CONTENTS

Title Page.............................................................................................................................i
Dedication............................................................................................................................ii
Acknowledgements.............................................................................................................iii
Table of Contents...............................................................................................................iv

1. Chapter 1: Introduction.....................................................................................................1
   1.1 Overview of Structured Interviews/Emotional Status Examination.......................3
   1.2 Overview of Dean-Woodcock Emotional Status Examination..............................11
   1.3 Overview of the Minnesota Multiphasic Personality Inventory............................13
   1.4 Diagnostic Utility and Rationale for Selecting the MMPI.......................................15
   1.5 Overall Rationale for Study.....................................................................................16

2. Chapter 2: Review of Literature.....................................................................................21
   2.1 Emotional Assessment............................................................................................22
   2.2 Minnesota Multiphasic Personality Inventory.........................................................23
   2.3 Clinical Interviews..................................................................................................40
   2.4 Emotional Status Examinations..............................................................................51
   2.5 Dean-Woodcock Emotional Status Examination..................................................60
   2.6 Summary................................................................................................................63

3. Chapter 3: Methods........................................................................................................65
   3.1 Participants..............................................................................................................65
   3.2 Instrumentation.......................................................................................................65
   3.3 Procedure................................................................................................................68
CHAPTER 1
INTRODUCTION

Over the past twenty-five years, the field of psychiatry has shifted to a medical/biological model and away from its roots in philosophy (see Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition-Text Revision; DSM-IV-TR; American Psychological Association, 2000). In fact, a number of researchers have argued in favor of biological based disorders with the majority of psychiatric symptoms and their interaction with neurological elements (see for a review Noggle & Dean, 2012). There are a number of ways in which these disorders may interact or be synergistic elements. A number of instances have been reported in which patients with neurological disorders present with related psychiatric disorders. For example, affective disorders and anxiety are likely to co-occur with numerous neurological disorders (e.g., TIA, stroke, or dementia). When psychiatric disorders were philosophically based, there were symptoms that did not meet the criteria for a “functional disorder” and were often referred to a psychiatrist for a differential diagnosis of functional vs. organic disorder. Measures of psychiatric symptoms found in structured and unstructured interviews were found to offer data based predictors, yet while measures such as the MMPI offered greater reliability,
information about the patient was lost. The purpose of this study is to the degree to which a relatively newly-developed structured diagnostic interview, the *Dean-Woodcock Emotional Status Examination* (D-WESE; Dean & Woodcock, 2003), which measures neurologically-based psychiatric symptoms, provides unique information not provided by the *Minnesota Multiphasic Personality Inventory* (MMPI; Hathaway & McKinley, 1942), a widely-used objective psychiatric inventory. The D-WESE may provide important information regarding psychiatric diagnosis and treatment that reflects the current shift to the medical/biological model.

The MMPI and its successor (MMPI-2) and revisions continue to be the most widely used and internationally adapted instruments for psychiatric diagnosis (Butcher, Derksen, Sloore, & Srigatti, 2003). An objective measure of psychiatric symptoms, the MMPI was constructed of more than 500 true or false questions that best predicted ten psychiatric diagnoses. Lubin, Larsen, Matarazzo, and Seever (1985) reported that the MMPI is the most frequently used test in the psychiatric clinical setting and it has been used as a benchmark to assess the validity of other measures of psychiatric disorders (e.g., Huckaby, Kohler, Garner, & Steiner, 1998; Liskow, Penick, Powell, & Haefele, 1986, Stangl, Pfohl, Zimmerman, Bowers, & Corentahl, 1985, Svanum & McGrew, 1996). However, some limitations to the use of the MMPI in research and practice have been noted. When the MMPI scales were developed, Hathaway and McKinley (1942) presumed the clinical scales would measure specific psychiatric disorders. However, research has indicated that elevations on the clinical scales predict behavior correlates rather than specific disorders. In fact, once the items related to a disorder are combined to create standard scores, it becomes difficult to examine individual symptoms. Further,
the MMPI clinical scales are composed of items which, on the average, overlap 69% with items on one or more other scales, suggesting poor definition across different psychological constructs (Shure & Rogers, 1965). In addition to providing little information for the clinician to work with, the cost and time required for administration of the MMPI can be prohibitive in current practice.

The D-WESE is a relatively newly-developed standardized clinical interview which is made up of items that were based on the diagnostic criteria for psychiatric disorders in the DSM-IV-TR. The information measured by the D-WESE is similar to that of the MMPI; however, the D-WESE items are based on current accepted diagnostic criteria in the field of psychiatry. Additionally, the D-WESE is less costly and time-consuming than the MMPI. However, although the items represent the diagnostic categories of the DSM-IV-TR, questions remain regarding the instrument’s ability to provide valid diagnostic scales of psychiatric disorders based on groups of items.

This study used a canonical correlation to examine the shared and unshared variability between the D-WESE and the MMPI. When the MMPI was in its zenith, the main goal of psychiatry was to differentiate between organic (biological) and functional (situational) disorders. Current trends in psychiatry favor the use of measures of psychiatric symptoms that are behavioral in nature.

**Overview of Structured Interviews/Emotional Status Examinations**

As discussed, emotional and psychiatric symptoms clearly play a major role in the diagnosis and treatment of neurological disorders. Indeed, research over the last thirty years has shown extensive comorbidity among neurological and psychiatric symptoms. Clinical diagnostic interviews have become an essential component of assessment,
diagnosis and treatment of emotional and psychiatric disorders. There are significant differences between structured and unstructured interviews. The Dean-Woodcock Structured Interview (D-WSI) is a structured diagnostic interview, which allows symptoms to be rated (e.g., on a 1-10 scale) and provides clinicians with the opportunity to elicit information regarding symptom occurrence, frequency, duration, severity, and time of onset that is not provided in objective measures of emotional functioning such as the MMPI.

Historically, unstructured and semi-structured interviews were used in which the interviewer determined the topic, wording, and order of questions to be asked. Unstructured and semi-structured approaches, however, have demonstrated poorer reliability and hence poor validity. For example, Ward, Beck, Mendelson, Mock, and Erbaugh (1962) found that clinicians who utilized unstructured interviews demonstrated considerable variability in diagnosis. According to this study, the majority of diagnostic variance (62.5%) resulted from discrepancies in the information clinicians deemed significant. Variations among clinicians in what questions were asked and how the resulting information was organized accounted for 32.5% of variance. Variations in patient response only accounted for 5.0% of variability (Ward, et al., 1962). In other words, 95% of the variability in clinician diagnosis was related to the utilization of unstructured interview techniques. Structured interviews, such as the Dean-Woodcock Emotional Status Exam (D-WESE), provide the opportunity to obtain behavior-based clinical information, and aid in the development of reliability and validity (Rogers, 2001).

Emotional status examinations have often been a major component of clinical interviews. The main purpose of an emotional status examination was to determine
whether or not the individual was experiencing clinical emotional difficulties. Emotional status examinations involved the use of particular sets of symptoms agreed upon by most clinicians as diagnoses that face validated specific a priori assumptions (Schwartz, 1989).

**History of clinical interviews.** Clinical interviews have a prominent place in the history of psychological assessment (Kamphaus & Frick, 2002). The face-to-face verbal dialogue between clinician and patient is the prototypical format for diagnostic interviews. Face-to-face dialogue between the patient and the therapist provides information that is not offered in a standardized test. Using an unstructured interview, the clinician determined what questions would be asked, how the questions would be framed, what follow-up questions were asked, and what were reasonable responses from the patient (Kamphaus & Frick, 2002). This approach for many experienced clinicians is found to be ideal as they ready themselves for specific therapy. Kamphaus & Frick (2002) reported that unstructured interviews rely heavily on the individual clinician’s orientation and expertise.

Early in the 20th century the approach to clinical assessment was rooted in the medical model. Detailed information regarding behaviors and symptoms was gathered in an effort to determine specific syndromes and diseases. Adolf Meyer (in Kaplan, Freeman, & Sadock, 1980) proposed a psychobiological approach, in which it was deemed important to determine biographical information of the client that included biological, historical, psychological, and social factors on current behavior. Engel (1977) expanded this approach into a biopsychosocial model. The data collected based on this model were used to reach a diagnosis. However, this approach lacked a reliable diagnostic system based on well established, mutually agreed upon criteria to diagnose
specific individual disorders (Shea, 1998). It should be recalled that psychiatric studies came from a history of notions of philosophy. Thus, a top down approach to study and evaluation of the patient was standard. In contrast, neurology, while working with rather small sample sizes, found its beginnings in the late 1800’s in biology.

Sigmund Freud’s psychoanalytic theory of psychiatric functioning and diagnosis of disorders moved from an emphasis on diagnosis, unlike other medical professionals. Freud believed emotional disturbances were less due to biological factors, establishing the notion of psychosexual developmental stages (Burger, 1990). The development of ego psychology and the investigation of defense mechanisms by theorists such as Heinz Hartman and Anna Freud shifted Sigmund Freud’s emphasis toward an understanding of how patients’ defenses were manifesting in the context of the interview itself (Burger, 1990). Psychoanalytic interviewing processes generally resembled an unstructured interview format.

In the early 1950’s, Harry Stack Sullivan established the significance of the interpersonal matrix as one of the major areas through which to understand the interview process (Sullivan, 1970). He stressed the importance of viewing the interview as a sociological phenomenon in which the client and clinician form a unique and dynamic dyad, with the behavior of each affecting the other. However, he argued that the client’s information is never objective and stated that subjective data are not valid due to the interference of clinician inference. “Inference,” according to Sullivan (1970), is one of the major problems in the development of practical psychiatric interviews. Indeed the act of measurement may change the outcome. In fact, the interview and treatment process were driven by the clinician’s theoretical orientation, which lacked quantitative
substantiation. Around this time, Karl Jaspers and Menard Boss emphasized the phenomenological approach to interviewing, which involved developing an understanding of the exact ways in which the patient experienced “being in the world” (Hall & Lindzey, 1978). This technique provided an opportunity for clinicians to understand clients’ symptoms, feelings, perceptions, and opinions as they related to their surroundings.

The non-objective fields of theory of psychology and counseling had the next impact on clinical interviewing (Shea, 1998). Indeed, these approaches emphasized the importance of utilizing the interview as a means by which the clinician could empathize with the client. Carl Rogers, for example, emphasized the “client-centered” approach, by which clinicians utilized empathic techniques to gain information (Rogers, 1951). Thus, clinical interviews would differ and offer few opportunities for comparison, limiting reliability and validity. However, as the field began to agree on specific disorders rather than complete idiosyncratic psychiatric features the notion of a structured interview became more and more reasonable and consistent with the science of the day.

The shift toward the more objective approach to interviewing reflected the overall choice of items to a more biological direction. This change can be seen with each revision of the Diagnostic and Statistical Manual of Disorders. As treatment modalities were being increasingly determined by diagnosis, better diagnostic systems that would improve validity and reliability needed to be developed (Shea, 1998). The publication of the DSM-III (1980) provided a more standardized approach to diagnosis with disorders having symptoms or behavioral criteria. This objective/behavior approach became more defined in revisions that followed. The DSM’s diagnostic system was accepted by a
growing number of clinicians beginning to use a more biological approach (Othmer & Othmer, 1994) and was integrated in the assessment of a patient’s emotional status. The work of Samuel Guze and Donald Goodwin at Washington University School of Medicine provided strong arguments and data in favor of the scientific approach to psychiatry, providing criteria for diagnosis using a medical model (North & Yutzy, 2010).

**History of structured interviews.** As structured interviews evolved, standards, symptoms, and clinical features became more consistent. These standards included the use of standard questions asked in interviews when making clinical inquiries, the sequencing of the inquiries, and the quantification of responses (Rogers, 2001). “Standardized diagnostic interviews” originated with the development of mental status examinations (MSEs) in 1917 by American psychiatrist Adolf Meyer (Tilley & Hoffman, 1981). Comprehensive MSEs were designed to measure/screen both cognitive impairment and psychopathology, in addition to providing a template for diagnosis (Rogers, 2001). In general, measures set a level of within normal limits. Development of comprehensive MSEs was spurred on by discontent with unstructured interviews and difficulty determining thorough, accurate diagnoses. However, these early measures did not meet psychometric standards of today.

Two early comprehensive MSEs, the Mental Status Evaluation Record (MSER) and the Missouri Mental Status (MMS), were developed in conjunction with state mental health programs in an effort to standardize clinical data (Rogers, 2001). The MSER, developed by Spitzer and Endicott (biological psychologists at NYU) in 1970, was designed to provide systematic ratings of psychopathology including psychotic
symptoms, somatic functioning and concerns, and interpersonal characteristics. Other areas of focus included symptoms of mood disorders, a gross estimate of intellectual functioning, and overt expression of anger and prior suicide attempts. The advantages the MSER provided over previously utilized unstructured interviews include attempting to reduce information variance by specifying the clinical coverage needed for a comprehensive assessment of DSM Axis I symptoms and features, attempting to reduce criterion variance by defining key terms, offering clear examples, providing gradients, and attempting to empirically validate its conceptualization of psychopathology (Rogers, 2001).

In an example of the utilization of an integrated system of clinical measures, the Missouri mental health system developed the MMS beginning in 1966 as one of seven measures used to follow the inpatient and outpatient interventions of over a hundred-thousand patients (Sletten & Evenson, 1972). The MMS consisted of clinical ratings based on assessment of appearance, motor activity, speech, and interview behavior, as well as sections on mood and affect, disturbances in thinking, and cognitive abilities (Rogers, 2001). Contributions of the MMS above and beyond those of the MSER include setting standards for the utilization of cross-validation with large diverse samples and evaluating a broad array of relevant clinical variables via sophisticated multivariate methods to provide validity to diagnostic decision making (Rogers, 2001).

Comprehensive MSEs such as the MSER and the MMS laid the groundwork for more current structured interviews, such as the Diagnostic Interview Schedule (DIS), the Schedule of Affective Disorders and Schizophrenia (SADS), the Structured Clinical Interview for DSM-IV Disorders (SCID), and multiple others (Rogers, 2001). The
continued development and use of structured interviews demonstrates clinicians’ prolonged commitment to reduce unnecessary variability and improve diagnostic agreement, and, by extension, provide more adequate care.

**Diagnostic utility of clinical interviews/emotional status examinations.**

Current structured interviews cover a wide range of information and utilize DSM diagnostic criteria in assessing a client’s emotional status. The reliability and validity data of specific structured interviews, measuring important patient information, can be gathered in various ways. One way would be rater agreement of results and the degree to which they make predictions of other forms of validity.

Structured interviews have been compared with a number of more objective and normative-based psychological assessment instruments (e.g., Beck Anxiety Inventory, Beck Depression Inventory, Millon Clinical Multiaxial Inventory). However, few studies have compared emotional status examinations to the MMPI. Due to the vast amount of research supporting the psychometric properties of the MMPI, it is often used to evaluate the validity of structured interviews (Kendall & Norton-Ford, 1982). A number of studies comparing MMPI profiles to the diagnoses and information obtained by structured interviews found similar diagnoses (e.g., Liskow, Penick, Powell, & Haefele, 1986; Lucas, 1974; Newmark, Gentry, Simpson, & Jones, 1978; Newmark, Gentry, & Whitt, 1983; Oberstone & Sukoneck, 1976; Stangl, Pfohl, Zimmerman, Bowers, & Corentahl, 1985; Svanum & McGrew, 1996).

For example, Newmark, Gentry, Simpson, and Jones (1978) utilized the Mental Status Schedule (MMS), a semi-structured interview, to determine the diagnosis of schizophrenia for 360 patients upon hospital admission. The MMPI was then
administered. Examination of the MMPI profiles revealed that 72% of the males and 71% of the females endorsed items suggesting schizophrenic symptoms. Stangl and colleagues (1985) examined the validity of the Structured Interview for the DSM-III Personality Disorders (SIDP) as compared to the MMPI. Participants who were diagnosed with personality disorders according to the SIDP scored significantly higher than those without personality disorders on all of the MMPI clinical scales.

Research correlating the MMPI with unstructured interviews, however, has shown less validity (Stangl, Pfohl, Zimmerman, Bowers, & Corentahl, 1985). For this reason, the MMPI was utilized in this study in part to determine the validity of a relatively newly-developed structured diagnostic interview. A significant correlation between the MMPI and the D-WESE would offer support for the use of the D-WESE in measuring psychiatric symptoms.

**Overview of Dean-Woodcock Emotional Status Examination**

The Dean-Woodcock Emotional Status Examination (D-WESE) is a structured diagnostic interview system that is based on the criteria of the Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition – Text Revision (DSM-IV-TR; American Psychiatric Association, 2000). The D-WESE is a component of the Dean-Woodcock Neuropsychological Battery that is used to measure the intensity and severity of emotional and behavioral problems. Preliminary research on the D-WESE revealed eleven factors, each measuring symptoms of mental problems similar to those measured by the MMPI (e.g., depression, anxiety, energy level, thought problems) (Galloway-Sharp, 2004).
Development of the Dean-Woodcock Emotional Status Examination. Dean and his associates constructed the measure with the key elements of structured interviews (e.g., clear, straightforward questions) in mind (Galloway-Sharp, 2004). The 50 questions were based on DSM-IV diagnostic criteria and were divided into three sections: (1) subject’s identifying information, (2) information concerning psychiatric signs and symptoms, and (3) clinical observations of the subject’s functioning. Section two elicits client information regarding depression, mania, attention, obsessive-compulsive behavior, anxiety, somatization, aggression, and paranoia. The D-WESE assesses the presence and severity of psychiatric symptoms in a structured manner. When symptoms are endorsed, the severity, as reported by clients, is rated. Examiner observations supporting or refuting the information are gathered and documented in a structured manner and used to aid in diagnosis. Few emotional status examinations include this component (Galloway-Sharp, 2004).

Rationale for selecting the Dean-Woodcock Emotional Status Examination. Although there are many psychological structured interviews, the D-WESE was chosen because it is believed to be a well-constructed measure that is relatively new and specifically examines neurological disorders. The items evaluate both the presence and severity of symptoms in a structured manner. A benefit of the D-WESE is that all of the important components are built into the same structured interview that is organized along the DSM-IV diagnostic criteria. The results of the interview provide clear symptoms useful in diagnosis, treatment, and psychological theory. Additionally, the D-WESE provides information regarding quality of life and the ways in which symptoms may interact with treatment (Dean & Woodcock, 2003). Therefore, the examiner can gain a
comprehensive picture of what the client is experiencing without having to use multiple different instruments. In contrast, the clinical scales of the MMPI provide behavioral correlates related to symptom groups (Greene, 2000). Therefore, information regarding certain symptoms, their duration and severity, and their interaction with treatment is lost. Thus, the MMPI may offer useful information in the overall assessment of emotional functioning, but it lacks information useful in understanding patient functioning. Unlike the MMPI, which requires hours to complete, the D-WESE requires about 20 minutes to complete, allowing for a broad range of information to be covered in a short period of time (Dean & Woodcock, 2003). The questions are straightforward and clearly stated. Level of severity and clinical observations can be documented. Because the D-WESE is new to the field, research should involve finding similarities to psychiatric diagnostic features. Preliminary research has demonstrated that factors of the D-WESE as derived through exploratory factor analysis correlate significantly with the clinical scales of the MMPI, indicating they share similar variability (Galloway-Sharp, 2004).

**Overview of the Minnesota Multiphasic Personality Inventory**

In a full neuropsychological assessment, clinicians often use a clinical interview, an emotional and/or mental status examination, and standardized testing. Since its development over fifty years ago, the Minnesota Multiphasic Personality Inventory has become the most widely used clinical testing instrument in the United States. A portion of the instrument’s popularity may be attributed to the multitude of critical research studies that have been performed to measure reliability and validity concerns over the past fifty years. The MMPI is an empirically derived measure that yields clinical scales and profiles providing information that can be interpreted by clinicians and translated to
clients with ease (Strupp, 1990). The MMPI clinical scales were originally designed to provide direct diagnoses based on scale elevations, so that an elevation on the Depression scale was indicative of a depressive disorder. This was true of each of the clinical scales. However, research did not support the use of the MMPI in this way. Instead, the relationship between scale elevations and behavioral correlates provides information regarding the client’s emotional state. Behaviors correlated with the most prevalent mental problems, including depression, anxiety, mania, thought disturbances, and hysteria are measured by the MMPI.

The MMPI is a self-report measure that assesses the patient’s behavior, attitudes, thought patterns, and strengths (Lanyon & Goodstein, 1972). It provides an objective measure of abnormal behavior and was designed to aid in diagnosing patients with mental illnesses in order to differentiate their illnesses between neurotic and psychotic disorders. In 1989, the MMPI was restandardized to provide norms for the inventory. Development of a national normative sample, provide appropriate representation of ethnic minorities, and update item content where needed, resulting in the MMPI-2 (Buther et al., 1989). However, the items on the validity and clinical scales of the MMPI were essentially unchanged on the MMPI-2 except for the elimination of thirteen items based on item content and the rewording of 68 items (Greene, 2000).

The MMPI has been used across a number of different settings, including employment agencies, university counseling centers, mental health clinics, schools, and industry. It has been used for assessment, screening, and research. It has become the assessment by which other inventories measure their utility (Kendall & Norton-Ford, 1982).
Diagnostic Utility and Rationale for Selecting the MMPI

The MMPI and MMPI-2 provide information regarding symptomology which aids examiners in reaching appropriate diagnoses. It assesses depression, anxiety, mental confusion, paranoia, mania, denial, social maladjustment, social introversion, and personal interests. Research has been conducted to support the overall predictive validity of the inventory as well as the accuracy of individual scales to measure the condition they purport to measure (Butcher & Owen, 1978).

Dahlstrom, Welsh, and Dahlstrom (1975) included almost 6,000 citations on the clinical and research applications of the MMPI. Overall, research has found the MMPI to be a reliable and valid measure (Alpert, 1999; Persinger & Tiller, 2002; Svanum & McGrew, 1996). The profiles obtained from the MMPI have been compared to results from similar measures of psychopathology (e.g., Millon Clinical Multiaxial Inventory, Rorschach). In general the MMPI became the “gold standard” in objective measures of psychopathology (Ganellen, 1996; Patrick, 1988; Pop-Jordanova, 2000; Zarella, Schuerger, & Ritz, 1990).

Criticisms of the MMPI include an overlap and redundancy in items and scales, outdated language, and the test’s length. The MMPI-2 was developed in an effort to address these concerns and to restandardize the MMPI. However, according to Dean (1986), the MMPI rather than the MMPI-2 was frequently utilized in neuropsychology due to the large data set available for comparative research. Also, due to the fact that only 13 MMPI questions were deleted from the standard Validity and Clinical scales, the decision to use one inventory over the other is relatively moot (Greene, 2000). Graham
et al. (1991) found a moderately high level of agreement between MMPI and MMPI-2 code types among well-defined profiles.

In a sample of psychiatric outpatients, the MMPI and MMPI-2 showed a modest convergence with provisional diagnoses made in five diagnostic categories (39% overall agreement) (Morrison, et al., 1995). The two inventories were equally able to identify psychopathology as defined by DSM diagnostic criteria. Overall, the predictive utility of the MMPI and the MMPI-2 have been found to be similar in clinical cases (Graham, Timbrook, Ben-Porath, & Butcher, 1991).

Many present clinicians may argue against the appropriateness of administering the MMPI in clinical practice. However, the purpose of the present study is research-based, to measure the amount and type of information the MMPI provides as compared to the D-WESE. Therefore, due to the MMPI’s rich, extensive literature, its large data set, and the similarities with the MMPI-2, the MMPI was used in this study. Additionally, noted neuropsychologist Ralph Reitan used the MMPI in research addressing emotional disturbances and psychological deficits as recently as 1997, indicating its continued importance in the field of neuropsychological research (Reitan & Wolfson).

**Overall Rationale for Study**

Measures of emotional function have long been seen as important in neuropsychological assessment. Reitan (1955) selected the MMPI as a measure of emotional function due to its objective approach. However, despite the authors’ original intent, derived clinical scales do not equate to psychiatric disorders. Instead, items measured by the MMPI are interpreted as they relate to the instrument’s constructs and the behaviors associated with those constructs. In an attempt to provide behavior
profiles, information on specific symptoms is lost. In fact, MMPI items identified as “obvious” were found to be better predictors of pathology than “subtle” items, suggesting some data measured by the MMPI is superfluous (Duff, 1965). In the past, clinical interviews were psychometrically unstable. Recently, however, structured interviews such as the D-WESE provide objective, reliable, and valid measures of emotional function.

Best practice in personality assessment involves using a number of sources to compile information that provides clinical and diagnostic impressions of individual clients (Nevid, Rathus, & Greene, 2011). The MMPI and the D-WESE are both essentially self-report measures; however, the MMPI has been highly researched and standardized. Information regarding underreporting and over-reporting of symptoms as well as response consistency and the potential for malingering are all present on the MMPI; however, evidence suggests clients may be able to produce profiles that are interpreted as valid, while still responding to test items in a manner that is not indicative of their personalities (Greene, 2000). The D-WESE allows for clinical observation to take place during the administration of the assessment. In mental health counseling, objective personality assessment is not always utilized in the process of therapy, as client problems may be seen as situational or diagnosable by different means. The MMPI is costly to administer and requires a great deal of patients’ time and energy. Clinical interviews are utilized despite the fact that the patient may have taken an MMPI to gather client information early in the course of counseling. Because insurance may only cover therapy for certain periods of time and for clients with certain diagnoses, the clinical interview becomes an important diagnostic and treatment tool.
The MMPI may offer a benchmark useful in overall assessment of emotional functioning, providing clinical validity to the diagnostic process. Additionally, the MMPI has been used in pretreatment planning, treatment progress, and posttreatment evaluation to determine treatment effectiveness and patient change (Butcher, 1990). However, the process of treatment requires more than patient diagnosis and evaluation of progress. Clinical interviews provide background information and useful insight in establishing diagnosis and treatment. This study will analyze the extent to which the two measures glean similar and different information. The amount of shared variability as demonstrated by the strength of the correlation between the two measures will provide information regarding the utility of the D-WESE in prediction and guidance of clinical treatment. Unshared variability will suggest the extent to which the MMPI and D-WESE measure different types of information, which may or may not be useful in psychiatric diagnosis and treatment. The purpose of this study is (1) to quantify the amount of shared variance between the MMPI scales and the D-WESE factors and (2) to calculate the amount of variance that could not be explained by the other measure.

A preliminary factor analysis of the D-WESE items yielded an eleven-factor model. However, subsequent research supported a different eleven-factor model based on confirmatory factor analysis and the theoretical constructs used in the development of the D-WESE. Based on the results of the confirmatory factor analysis, several predictions can be made regarding the relationship between the clinical scales of the MMPI and the D-WESE factors. For example, many of the MMPI clinical scales and D-WESE factors purport to measure similar emotional constructs. The MMPI clinical scale names were created by the items’ ability to distinguish individuals with and without
psychiatric disorders while the D-WESE items were selected based on their inclusion in DSM-IV-TR diagnostic criteria. Both reflect the major diagnostic categories relevant at the time the measures were developed. The MMPI, for example, included a scale designed to identify homosexuality, which was considered a psychiatric disorder at the time. This scale (Scale 5: Masculinity-Femininity) is currently used to determine masculine vs. feminine interests and is not given much credence in research and practice (Greene, 2000). Therefore, it is unlikely that Scale 5 will correlate significantly with the factors of the D-WESE.

Also, the D-WESE measures Intellectual/Executive Functioning and Attention, the research for which has increased markedly over recent years. Executive functioning involves the cognitive processes that regulate, control, and manage other cognitive processes, such as planning, working memory, attention, problem solving, verbal reasoning, inhibition, mental flexibility, task switching (Monsell, 2003). The assessment of executive functioning is an important component of neuropsychological evaluations. The MMPI does not include scales that measure Executive Functioning and Attention. Therefore, it is unlikely that the Intellectual/Executive Functioning and Attention factors will correlate significantly with the MMPI. Based on the empirical and theoretical selection of the items, the remaining MMPI scales and factor scores measure psychological constructs that are likely to correlate with each other in the following manner, with the MMPI scale listed first followed by corresponding D-WESE factor scores:

Scale 1 (Hypochondriasis): Somatic Symptoms

Scale 2 (Depression): Depressive Features, Energy Level, Appetite/Sleep
Scale 3 (Hysteria): Somatic Symptoms

Scale 4 (Psychopathic Deviate): Asocial Behaviors, Rebelliousness

Scale 6 (Paranoia): Anxiety Symptoms, Unstable Mood/Psychotic Symptoms

Scale 7 (Psychasthenia): Anxiety Symptoms, Posttraumatic Symptoms

Scale 8 (Schizophrenia): Unstable Mood/Psychotic Symptoms

Scale 9 (Hypomania): Depressive Features, Unstable Mood/Psychotic Symptoms, Energy Level, Appetite/Sleep

Scale 0 (Social Introversion): Anxiety Symptoms

As stated, these relations reflect the theoretical and empirical methods by which the scales of psychiatric constructs were derived. For example, mood symptoms as measured by the MMPI should correlate with mood symptoms as measured by the D-WESE. However, if the factors do not correlate as suggested, results may suggest the MMPI and D-WESE measure the same psychological constructs in different ways that reflect the changes in the way psychiatric information is gathered.
Standardized assessment of emotional functioning was encouraged to aid in psychiatric screening and treatment following World War I (Greene, 2000). Early personality inventories were constructed on a philosophical basis and varied in reliability and validity (Greene, 2000). Using a negative/positive method of choosing items, Hathaway and McKinley (1940) developed the Minnesota Multiphasic Personality Inventory (MMPI), one of the few empirically derived measures of emotional disorders. Composed of some 500 true/false items, the MMPI became one of the most widely used and researched objective emotional inventories. However, Cella and Perry (1986) argued that completing questionnaires of such length can be “very taxing” for patients already acutely stressed (p. 827). Many of the MMPI items are known as “subtle”, which means these items did not relate to the clinical scales in an obvious way. The MMPI continues to be of interest in assessing the ability of less time-consuming approaches to diagnosis. Although the MMPI has been revised, the vast majority of items remained the same, the standardization of the MMPI-2 was based on “normals” and the available research on the MMPI is overwhelming. This chapter examines the shared variability of the MMPI
and structured clinical interviews. More specifically, it will consider the shared variability of MMPI and a newly developed emotional status examination, the Dean-Woodcock Emotional Status Examination.

**Emotional Assessment**

Two forms of emotional assessment involve diagnostic inventories and clinical interviews. Diagnostic inventories such as the MMPI are usually completed by the patient using a pencil and paper format, as well as newer computerized approach. The questions asked address the various behaviors and feelings associated with emotional disorders, which follow diagnostic criteria set by the DSM (American Psychological Association, 2000). The patient is required to acknowledge or deny the symptom/experience. Some inventories also ask the patient to indicate the severity and/or frequency of the symptoms reported. Diagnostic inventories vary in their breadth. Some focus on single emotional disorders such as anxiety or depression, while others attempt to investigate several emotional disorders. Strengths of diagnostic inventories include greater standardization without the need to establish rapport, and the possibility of obtaining information from clients who have difficulty expressing themselves and/or articulating what they are experiencing. These individuals may be willing to admit symptoms and behaviors they are experiencing in a pencil and paper format versus speaking face-to-face with a clinician.

Interviews have long been used to gather clinical information. Conducted in a face-to-face format, the clinician collects information which fits with the clinician’s hypothesis. Topics similar to those addressed in diagnostic inventories are assessed in a conversation-style format, with clinicians recording patient responses. The rigidity and
structure of clinical interviews vary, as do the psychometric standards. Clinical interviews more often explore the symptomology of a wide range of emotional disorders, rather than focusing on single disorders. Strengths of clinical interviews include the opportunity to go beyond the standardized questions and to clarify and observe. Clinical interviews also allow patients to elaborate on their emotional status. Additionally, the influence of demographic variables such as ethnicity and race can be more easily considered.

Clinicians obtain the most valid diagnoses by combining information from both diagnostic inventories and clinical interviews. The reliability of measures increases with structure. Therefore it is important to explore the research and practical utility of both types of assessments, as well as the relationship between the two.

**Minnesota Multiphasic Personality Inventory**

The MMPI was an early attempt to standardize elements of emotional assessment, which would correspond with diagnostic symptoms of neurotic and psychotic disorders (Greene, 2000; Lubin, Larsen, Matarazzo, & Seever, 1985). Although the authors originally hoped that diagnosis would correspond to specific individual disorders (i.e., schizophrenia) the measure of emotional functioning was found to relate to a combination of different symptomology (Greene, 2000). In sum, its effectiveness has been in predicting the validity of structured interviews/emotional status examinations (e.g., Huckaby, Kohler, Garner, & Steiner, 1998; Liskow, Penick, Powell, & Haefele, 1986; Stangl, Pfohl, Zimmerman, Bowers, & Corentahl, 1985; Svanum & McGrew, 1996).

Hathaway and McKinley (1940) developed the MMPI using a strict empirical approach, selecting items that address a wide sample of behavior for a wide range of
psychiatric disorders. These items consist of short statements intended to assess psychopathology. For example, “The future seems hopeless to me.” “I have few or no pains.” The basic purpose of the test was to differentiate between various types of mental patients, as well as to distinguish between mental patients and “normal” people.

Hathaway and McKinley (1940) accumulated a large pool of items from which various scales were constructed, in the hopes of involving a greater variety of valid personality descriptions than was previously available. The items asked on the MMPI are designed to evaluate the thoughts, emotions, attitudes, and behavioral traits that comprise personality. The results of the test reflect an individual’s personality strengths and weaknesses, and may identify certain disturbances of personality (psychopathologies) or mental deficits caused by neurological problems (Graham, 1999). The MMPI seeks to identify individuals with various psychological disorders such as depression and schizophrenia.

The primary normative group for the MMPI included individuals ranging in age from 16 to 55. In current practice, it is suggested that the adolescent version of the MMPI (MMPI-A) be used for individuals between the ages of 14 and 18. However, Williams and Butcher (1989) provided support for the use of MMPI scale descriptors from studies based on adults as the interpretive strategy for a large adolescent sample ranging in age from 12-18, providing support for the use of the MMPI in research with a younger sample.

The issue of whether the items and patient norms for the MMPI developed in the early 1940s are appropriate for contemporary use has been raised repeatedly and debated widely (Butcher, 1972; Colligan, Osborne, Swenson, & Offord, 1983; Faschinbauer,
Due to the fact that the typical individual in the original Minnesota normative group was approximately thirty-five years old, married, living in a small town or rural area, had eight years of general education, and worked at a skilled or semi-skilled trade (or was married to a man with such an occupation level), it seems apparent that there have been numerous changes in our society over the ensuing six decades (Dahlstrom, Welsh, & Dahlstrom, 1972). Pancoast and Archer (1989) reviewed the existing literature on the performance of normal individuals on the MMPI to assess the adequacy of the norms based on the original Minnesota normative group. They found that the scores of normal individuals across five decades differed only slightly from the original Minnesota normative group on the standard validity and clinical scales on the MMPI.

Greene (1991) examined the changes in the validity and clinical scales on the MMPI within four frequently occurring codetypes (i.e., combinations of scale scores) in samples of psychiatric patients over a span of 40 years. He found that the MMPI scale scores of psychiatric patients had been very stable over this time span. The finding that normal individuals and psychiatric patients have shown only minimal changes on the standard validity and clinical scales on the MMPI across 40 years would suggest that the inventory may not be as outdated as some have proposed (Greene, 2000).

The *Minnesota Multiphasic Personality Inventory-2* (MMPI-2) was developed in 1989 to address the criticisms of the original version. The MMPI-2 provided current norms, developed a nationally representative and larger “normal” sample, provided appropriate representation of minority groups, and updated item content where needed. The MMPI-2 represents the re-standardization of the MMPI. However, both versions
have been shown to produce similar results (e.g., Graham, Timbrook, Ben-Porath, & Butcher, 1991; Morrison et al., 1995). Also, the problem of time for administration continues with the MMPI-2 (approximately three hours). See John Graham’s MMPI-2: Assessing Personality and Psychopathology for additional information.

**Structure of the MMPI.** The MMPI consists of 566 true/false items designed to assess the emotional elements of adolescents and adults. The examinee’s scores provide information on scales in the following categories: validity, clinical, and research. In answering the items of the inventory, examinees report information regarding the amount of psychopathology they are feeling as well as their willingness to share such information.

**Validity scales.** When developing the MMPI, Hathaway and McKinley understood the importance of assessing whether or not the client has responded appropriately to the inventory. Therefore, the MMPI was one of the first tests of emotion to offer a means of directly assessing a client’s test-taking attitude, providing examiners with empirical means to address this variable (Greene, 2000). The MMPI contains three validity scales, Lie (L), Infrequency (F), and Defensiveness (K). The validity scales measure both the validity of the MMPI in terms of how well the items measure what they purport to measure and the validity of the client’s responses. More specifically, the validity scales provided data regarding the client’s tendency to distort symptom information and/or provide an accurate and consistent self-description (Greene, 2000).

Multiple factors contribute to clients providing inaccurate descriptions of themselves. Invalid scores on the validity scales may indicate true psychopathology, a personality disorder, or a client’s cry for help. Alternately, clients may respond to items
inappropriately due to ambiguous wording, which can cause different interpretations of the same item. Clients may also vary in the level of self-awareness they possess and the extent to which they are willing to share information (Greene, 2000).

Meehl and Hathway (1946) believed in the importance of assessing two dichotomous categories of test-taking attitudes: defensiveness (underreporting) and plus-getting (overreporting). Clients who underreport their symptoms endorse a lesser number of symptoms than they are actually experiencing, while clients who overreport their symptoms endorse a greater number of symptoms than they are actually experiencing. The validity scales were designed to measure the extent to which the client over- or underreports his/her symptomology.

To assess over/underreporting, Meehl and Hathaway (1946) utilized three approaches. First, they gave the client an opportunity to distort the responses in a specific way and observed the extent to which the client did so (Greene, 2000). For example, they repeated items throughout the inventory to assess the consistency in client responses. However, this method for assessing over/underreporting was ultimately rejected. Next, Meehl and Hathaway provided an opportunity for the client to answer favorably when a favorable response would almost certainly be untrue (Greene, 2000). This involved developing a list of extremely desirable but very rare human qualities. Additionally, Meehl and Hathaway provided items endorsed with relative infrequency by the normative group, allowing clients to answer. Finally, Meehl and Hathaway determined items by which individuals known to have psychopathology and who were hospitalized who obtained normal profiles from normal individuals who obtained
elevated profiles (Greene, 2000). These types of controls are not available with the structured interview.

The validity scales were established by combining particular items contained within the MMPI clinical scales (Woychyshyn, McElheran, & Romney, 1992). The $L$ (Lie) scale consists of items that, if the patient were to endorse in large numbers, there would be a high probability that he/she was being dishonest (Kroger & Turnbull, 1975). The scale is used predominantly to measure the extent to which clients claim social virtue, as well as their openness and willingness to share emotions and behavior. The $L$ scale provides information regarding the extent to which a patient may be underreporting psychopathology.

The $F$ (Infrequency) scale was developed according to a variant of this approach. This scale identifies individuals who have endorsed an abnormally high number of items that are infrequently admitted indicating psychological distress. Clients with high scores on the $F$ scale may be intentionally faking pathology, which is considered to be overreporting psychopathology. This scale is made of items that are infrequently endorsed by the normative population (Greene, 2000).

The $K$ (Defensiveness) scale measures the amount of good feelings and resources the patient reports having (Greene, 2000). It also quantifies the extent to which the client attempts to answer questions in a socially desirable fashion. Elevated scores on the $K$ scale may be indicative of underreporting psychopathology.

**Clinical scales.** The clinical scales provide information addressing the DSM diagnostic criteria. The questions assist in the diagnosis of Axis I and Axis II disorders. The MMPI originally consisted of eight clinical scales (1-4, 6-9). In 1946, scales 5 and 0
were added providing ten clinical scales. The names of the scales and their abbreviations follow (Greene, 2000).

1 Hs  Hypochondriasis  
2 D   Depression       
3 Hy  Conversion Hysteria  
4 Pd  Psychopathic Deviate  
5 Mf  Masculinity-Femininity  
6 Pa  Paranoia  
7 Pt  Psychasthenia  
8 Sc  Schizophrenia  
9 Ma  Hypomania  
0 Si  Social Introversion  

Scale 1, Hypochondriasis, measures a wide variety of vague bodily complaints. These complaints, which persist despite negative medical tests, primarily focus on the abdomen and back. Scale 1 is designed to assess neurotic concern over bodily functioning and whether such concern is used to manipulate others. A person who is actually physically ill would typically only obtain a moderate elevation (T=58-64; clinically significant elevation=70) because he/she only endorses items that are legitimate physical complaints (Greene, 2000).

Scale 2 measures symptomatic depression, which is characterized by a generally depressed mood, dissatisfaction with life, and no hope for the future (Hathaway & McKinley, 1940). The content of this scale deals with a lack of interest in activities reflected by apathy, physical symptoms, excessive sensitivity, and withdrawn personality.
This scale is most frequently the highest elevation in psychiatric profiles (Duckworth & Anderson, 1986).

Scale 3, Conversion Hysteria, determines the amount of denial a person exhibits regarding any conflict he/she is experiencing. Items measure symptoms typically in the head, arms, and legs as well as information regarding whether the client considers himself or herself to be well socialized and well adjusted. In individuals with histrionic dynamics, these two categories are closely associated. High scores on Scale 3 indicate clients who are self-centered, immature, and infantile (Greene, 2000).

Scale 4 measures general social maladjustment and the absence of strongly pleasant experiences (McKinley & Hathaway, 1944). Concerns about family and authority figures in general, self- and social alienation, boredom, denial of social shyness, and the assertion of social poise and confidence are measure on this scale. Individuals with high scores on scale 4 are often described in unfavorable terms such as angry, impulsive, emotionally shallow, and unpredictable (Greene, 2000).

Scale 5, Masculinity-Femininity, measures interests in vocations and hobbies, aesthetic preferences, activity-passivity, and personal sensitivity. Responses to these items are scored as deviant when they reflect femininity in men and masculinity in women (Greene, 2000). Duckworth and Anderson (1986) describe this scale as the most misunderstood of all the Clinical scales for the following reasons: (1) the assumption that the scale can determine one’s level of masculinity or femininity, (2) the implication that the scale can detect homosexuality, and (3) separate interpretations of Scale 5 for males and for females. Interpretation of the scale is based on the extent to which individuals identify with their stereotyped gender roles.
Scale 6 is made up of items assessing interpersonal sensitivity, moral self-righteousness, and suspiciousness. The content of some items is clearly psychotic, acknowledging the existence of delusions and paranoid thought processes. High scorers on scale 6 are generally described as being suspicious, hostile, guarded, overly sensitive, argumentative, and prone to blame others. Clients who score in the normal range may not have any paranoid symptomatology or may have deeply ingrained paranoid symptomatology and are avoiding obvious items on Scale 6 (Greene, 2000).

An inability to resist specific actions or thoughts regardless of their maladaptive nature originally characterized as the neurotic syndrome of psychasthenia is measured by Scale 7. Scale 7 also assesses abnormal fears, self-criticism, difficulties in concentration, and guilt feelings. The anxiety assessed by Scale 7 is of a long-term nature or trait anxiety, although the scale is somewhat responsive to situational stress as well (Greene, 2000). High scores on Scale 7 indicate anxiety, tension, indecisiveness, and an inability to concentrate. Individuals with elevations on Scale 7 frequently display obsessive thoughts and ruminations, self-doubt, and associated depressive features.

Scale 8, Schizophrenia, assesses a wide variety of content areas, including bizarre thought processes and peculiar perceptions, social alienation, poor familial relationships, difficulties in concentration and impulse control, lack of deep interests, disturbing questions of self-worth and self-identity, and sexual difficulties. Individuals with elevated scores on Scale 8 are described as cold, apathetic, alienated, misunderstood, and having difficulties in thinking and communication, which may reflect an actual psychotic thought disorder. Chronic schizophrenics who have adjusted to their psychotic process may score within the normal range on Scale 8, whereas “normal” individuals with
slightly elevated scores on Scale 8 may be described as self-dissatisfied, irritable, having wide interests, and immature (Greene, 2000).

Scale 9 measures the milder degrees of manic excitement, characterized by an elated but unstable mood, psychomotor excitement, and flight of ideas. The items range over a wide variety of content areas including behavioral and cognitive overactivity, grandiosity, egocentricity, and irritability. Scale 9 can be conceptualized as providing energy to activate the qualities identified by other elevated clinical scales (Greene, 2000). High scores on Scale 9 indicate impulsive, competitive, talkative, narcissistic, amoral, extroverted individuals who are superficial in social relationships.

Social introversion-extroversion is assessed by Scale 0, with high scale scores reflective social introversion. The social introvert is described as uncomfortable in social interactions and who typically withdraws from such interactions when possible. These individuals may have limited social skills or simply prefer to be alone or with a small group of friends. The social extrovert is socially outgoing, gregarious, and seeks social interactions. Item content on Scale 0 reflects personal discomfort in social situations, isolation, general maladjustment, and self-deprecation (Greene, 2000).

**Construction of the clinical scales.** Scale 1, Hypochondriasis, was the first scale to be constructed (McKinley & Hathaway, 1940). Hypochondriacs were one of the largest groups of patients available to McKinley and Hathaway and hypochondriasis is one of the simpler, more definite diagnostic categories, (Greene, 2000). The first step in developing Scale 1 was to select an appropriate criterion group. Diagnostic classifications were used, as the intent of MMPI was to aid in differential diagnosis (Greene, 2000). McKinley and Hathaway chose 50 cases of pure, uncomplicated
hypochondriacs as their criterion group. They then chose 724 individuals who were friends or relatives of patients in the University Hospitals in Minneapolis who were not currently receiving treatment from a physician to serve as the normative group. Four additional normative groups were utilized to assess the influence of “nuisance” variables such as age, socioeconomic status, or education and to include groups of individuals hospitalized for some form of physical disease and individuals who were patients in the psychopathic unit of the University Hospitals, regardless of diagnosis (Greene, 2000).

Once the criterion group and the normative groups were established, the process of item selection began. For the criterion group and each of the normative groups, the frequency of “true” and “false” responses was calculated for each item. An item was considered significant and was tentatively selected for the scale if the difference in frequency of response between the criterion group and the normative groups was at least twice the standard error of the proportions of true/false responses of the two groups being compared (Greene, 2000). Some of the items meeting significance were eliminated due to a low percentage of the criterion group’s response, item bias on variables such as marital status or socioeconomic status, and items that, upon inspection, were not germane to the construct of hypochondriasis. The prospective items were then weighed and combined into one scale (Greene, 2000).

Scales 2, 3, 4, 6, 7, 8, and 9 were constructed empirically in a similar fashion to that of Scale 1, with the same five normative groups being utilized for each scale. The other scales included the following criterion groups:

Scale 2 – 50 patients who represented relatively uncomplicated cases of the depressed phase of manic depressive psychosis.
Scale 3 – 50 patients with either a diagnosis of hysteria or identifiable histrionic personality components.

Scale 4 – young adults (ages 17 to 22 years) diagnosed as psychopathic personalities, asocial, and amoral types.

Scale 6 – an unspecified group of paranoid patients.

Scale 7 – 20 patients whose final diagnoses was psychasthenia.

Scale 8 – 50 patients diagnosed with assorted types of schizophrenia.

Scale 9 – 24 manic patients of mild or moderate severity.

Scales 5 and 0 were developed in a slightly different manner than the other Clinical Scales. Approximately 55 items, mostly related to sexual orientation, were added to the MMPI item pool after data had been collected for the original normative sample. Therefore the criterion group of male homosexuals who were used in developing Scale 5 could not be contrasted with the original normative group on these 55 items. Consequently, a normative group of 54 male soldiers was used for this scale, with items that distinguished them from the male homosexuals being included on Scale 5. In addition, items that differentiated men from women within the normative scale were included on the scale (Greene, 2000).

In 1946 Scale 0 was added to the MMPI (Drake, 1946), completing the standard MMPI clinical profile. Drake selected MMPI items that differentiated 50 college students who scored above the 65th percentile on the Minnesota T-S-E Inventory (Evans & McConnell, 1941) from 50 students who scored below the 35th percentile. Although the Minnesota T-S-E Inventory assesses introversion-extroversion in the areas of
thinking, social, and emotional, Drake limited his investigation to the social introversion-extroversion area.

**Interpretation of the clinical scales.** As previously mentioned, Hathaway and McKinley (1940) originally intended the clinical scales to be used to place patients into various diagnostic categories. For example, a person with an elevated score on Scale 2 (Depression) would be diagnosed as depressed. It was later discovered that this was a very limited approach to the interpretation of the clinical scales, leading to three major developments (Duckworth & Anderson, 1986). First, clinicians began to describe patients based on the behavioral interpretations of scale elevations, rather than by the name of the scale (e.g., a person with an elevated score on Scale 2 would be described as experiencing a lack of interest in activities, general apathy, sleep disturbances, excessive sensitivity, and lack of sociability). This interpretation style provided useful information for treatment (Duckworth & Anderson, 1986). Second, clinicians began to consider the varying scale elevations to differentiate the intensity of a behavior or thinking pattern (e.g., “blue” vs. severely depressed). Third, clinicians analyzed the whole profile rather than the few high points, adding subtlety and richness to the profile interpretation (Duckworth & Anderson, 1986).

**Research scales.** In addition to the validity and clinical scales, a number of research scales have been developed to measure a variety of areas such as alcoholism, ego strength, dominance, anxiety and status needs (Duckworth & Anderson, 1986). These include: First Factor (Conscious Anxiety), Second Factor (Conscious Repression), Ego Strength, Low Back Pain, Causality, Dependency, Dominance, Social Responsibility, Prejudice, Social Status, Control, MacAndrew Addiction Scale. The
research scales were not utilized in this study, and therefore will not be discussed in further detail in the scope of this paper.

**Reliability and validity of the MMPI.** The MMPI has been found to aid in the diagnosis of psychiatric patients (Meehl & Dahlstrom, 1960). It is widely used and frequently researched (Butcher, 1987, Svanum & McGrew, 1996). Overall, research has found the MMPI to be a reliable and valid measure (Alpert, 1999; Persinger & Tiller, 2002; Svanum & McGrew, 1996).

There have been a number of studies demonstrating the reliability of the MMPI. In general, this research provides solid evidence of replicable correlates of test scores to non-test behavior across a broad and diverse range, including such things as impulsiveness, depression, alcoholism, and success as a police officer (e.g., Beutler, Storm, Kirkish, Scogin, & Gains, 1985; Marks, Seeman, & Haller, 1974). The reliability of the MMPI has been demonstrated through test-retest agreement and in comparison to other instruments (e.g., Pancoast, Archer, & Gordon, 1988; Parker, Hanson, & Hunsley, 1988; Persinger & Tiller, 2002; Schuerger, Tait, & Tavernelli, 1982; Schuerger, Zarrella, Hotz, 1989). In general, studies have shown a statistically significant level of test-retest reliability over a period of two months to a year. Researchers have demonstrated that the MMPI has a statistically reliable degree of association with the DSM-based categories of psychopathology (Svanum & McGrew, 1996).

Alpert (1999) reported the MMPI is a valid, stable, and reliable instrument when it is used in the manner for which it was designed and validated. The validity of the MMPI has been demonstrated through a large number of studies (Dahlstrom, Welsh, & Dahlstrom, 1972). The concurrent validity of the MMPI in relation to DSM classification
system was measured by classifying gross diagnostic groups (normal, deviant, neurotic, psychotic, and sociopathic) by using regression techniques (Goldberg, 1972). Hit rates of over 90 were obtained for all diagnostic classification distinctions.

In a meta-analysis of 403 studies on the ability of the MMPI to identify psychiatric diagnoses as defined by the DSM, similar results to Goldberg’s research were found (Zaleski & Gottesman, 1991). At the level of group data, multivariate equations based on combinations of MMPI scales were relatively successful in classifying cases into broad diagnostic categories. Zaleski and Gottesman reported an overall reliability of .81 and almost 100% accuracy in discriminating between control and psychiatric groups in general (i.e., presence or absence of any disorder). MMPI discriminations among psychiatric categories, however, were less accurate. The strongest discrimination was between neurotic and psychotic disorders, whereas the weakest was between schizophrenia and affective disorders. When using single scale elevations to make a diagnosis, hit rates of only 53% to 64% were obtained in making the correct diagnosis (Goldberg, 1965; Loy, 1959; Meehl & Dahlstrom, 1960; Taulbee & Sisson, 1957). Therefore, researchers suggest it is most appropriate to interpret MMPI results based on a combination of the symptoms associated with the highest elevated clinical skills.

Svanum & McGrew (1996) compared a structured interview, the Diagnostic Interview Schedule (DIS), to the MMPI. The DIS strictly adheres to the DSM-III diagnostic criteria and was used to determine the presence of pathology and the specific disorder present. The MMPI was able to correctly identify 77% of the individuals who had not received a diagnosis according to the DIS and to correctly identify 54% of those who had received a diagnosis.
In addition to comparing the MMPI with DSM diagnostic criteria, the MMPI has been validated by comparison with other personality inventories (e.g., Millon Clinical Multiaxial Inventory, Rorschach). Patrick (1988) found that the MMPI was more effective than the Millon Clinical Multiaxial Inventory (MCMI) in identifying schizophrenic or psychotic patients. Ganellen (1996) compared the diagnostic efficiency of the MMPI, the Millon Clinical Multiaxial Inventory-II (MCMI-II), and the Rorschach for detecting depressive disorders. The three measures were comparable in their sensitivity in detecting depression. However, high scores on the MMPI and the MCMI-II tended to over identify patients as depressed.

In a group of 100 private-practice psychology patients, MMPI scale scores were used to predict MMCI scores (Zarrella, Schuerger, & Ritz, 1990). Scores on 19 of the 20 MCMI scales could be successfully predicted by the selected MMPI scales. Pop-Jordanova (2000) conducted a study comparing the MMPI, the Eysenck Personality Questionnaire, the General Anxiety Scale, and the Cornell Medical Index results in preadolescents with eating disorders. Similar results were found across these measures, especially in the area of anxiety.

Dunteman and Bailey (1967) researched the relationship between the MMPI and the Strong Vocational Interest Blank (SVIB). The results of the canonical correlation analysis indicated there is more of an overlap between inventoried personality and interest traits than prior results had demonstrated. The authors suggested the use of a canonical correlation, as opposed to factor analysis, seemed to bring out the similarities. Similarly, Vestre and Lorei (1967) used a canonical correlation analysis to compare clients’ MMPI scores with historical information gathered from the client’s nearest
relative. The results suggested a relationship between social history and symptomatic information gathered from the client.

**Criticisms of the MMPI.** The MMPI has been open to a number of criticisms. Perhaps the most frequently voiced concern is that some of the scales are highly correlated, indicating a considerable degree of overlap (Greene, 2000). Similarly, Block (1965) has shown that many of the “pathological” items are highly similar in nature, creating a redundancy in items as well as scales. Other common criticisms are that the test takes too long to administer, that it contains items about sex and religion that are offensive to many people, and that the wording is outdated (Greene, 2000).

In addition to the above-mentioned criticisms, the MMPI has not been completely successful in its original purpose, the classification of psychiatric patients (Lanyon & Goodstein, 1971). High scorers on scales often do not fit into the diagnostic category associated with the scale name, and large numbers of apparently normal people achieve high scores on the clinical scales (Greene, 2000). The latter is likely due to the fact that the MMPI was not designed for the assessment of normal personality. Therefore, neither the items nor the scales are optimal for this purpose (Lanyon & Goodstein, 1971). Some of the concerns raised regarding the MMPI led to examination of combinations or patterns of scores in order to make useful clinical and personality discriminations (Lanyon & Goodstein, 1971).

**Rationale for selection of the MMPI.** Despite the criticisms, the MMPI remains the most widely used personality inventory across a variety of clinical settings (Butcher & Rouse, 1996). The MMPI is a reliable and valid instrument that has been shown to aid in the diagnostic process (Dahlstrom, Welsh, & Dahlstrom, 1972; Persinger & Tiller,
2002). It has a large research base and norms have been established. The MMPI has long-standing credibility as the most respected and utilized personality test in the United States (Huckaby, Kohler, Garner, & Steiner, 1998; Keller & Piotrowski, 1989; Lubin, Larsen, & Matazarro, 1984). It has been referred to as “the gold standard of psychological assessment” (Kendall & Norton-Ford, 1982, p. 310). The content validity of the MMPI is thought to be high because it assesses pathology in conjunction with the DSM. It has also been used to measure concurrent validity for a number of other personality inventories (Ganellen, 1996; Patrick, 1988). Due to the relatively high validity of the MMPI, Sloore (1988) stated that it could, without any doubt, help clinicians make predictions about behavior.

Although the MMPI can contribute to the diagnosis of a client, the reliability and validity of clinical diagnoses increases as other sources of clinical information are involved in the decision-making process (Pancoast, Archer, & Gordon, 1988). Findings support the conclusion that diagnoses based exclusively on the MMPI are likely to be wrong more times than they are right in terms of congruence with psychiatric clinical diagnoses (Pancoast, Archer, & Gordon, 1988). Therefore it would be most effective if used in conjunction with other structured sources of information. Its use with structured interviews will be discussed further in a later section.

Clinical Interviews

Clinical interviews have a prominent place in the history of psychological assessment. Harry Stack Sullivan defines a psychiatric interview as a voluntary, primarily vocal communication between two people, one of whom is an expert and the other a client, working in a highly technical manner, with the purpose of determining the
characteristic patterns of living of the client and which of these patterns he experiences as particularly troublesome or valuable, the revelation of which brings some benefit to the client (Stack Sullivan, 1970). Interviews provide an opportunity to gain information directly from the patient about symptoms and their duration and severity, as well as to utilize clinical judgment/impressions regarding consistency between reported and observed symptoms.

The human contact that is achieved during an interview can have a positive effect over that of pencil and paper inventories. The nature of diagnostic interviews allow for the opportunity to build rapport, relieve patient distress, and observe behavior. Lanyon and Goodstein (1982) elaborated on reasons why clinical interviews are valuable assessment instruments. First, even highly structured interviews are more flexible than paper-and-pencil or computer assessments because they allow for the possibility of further examination of unclear responses. Second, the clinician is able to establish rapport with the patient, which is less likely to occur with pencil-and-paper or computer instruments. Third, patients in general, and confused or distressed patients in particular, are better able to relate to a clinician during an interview than when completing a written questionnaire. The interview provides opportunities for patients to express their concerns. This is particularly relevant for racial and ethnic minority persons who may feel that they will be misrepresented by traditional paper-and-pencil instruments (Sabnani & Ponterotto, 1992). The nature of the two-way communication allows for clarification of issues, providing more accurate assessment and promotion of patient commitment to and satisfaction with the diagnostic process (Hodges & Cools, 1990). Finally, clinical interviews permit a clinician to observe patients’ nonverbal reactions to various subjects.
during the interview. Clinicians who are sensitive to subtle patient behaviors may be more likely to pursue questioning related to areas that are particularly important, leading to more accurate diagnosis and promoting a perception that the clinician is interested in the patient and invested in thoroughly understanding his/her concerns.

The main disadvantage of the clinical interview compared with paper-and-pencil or computer instruments is its sources of error and subjectivity. Stack Sullivan believed client information is never objective and states that subjective data are not valid due to the interference of practitioner inference. Inference, according to Stack Sullivan is one of the major problems in the development of practical psychiatric interviews (Stack Sullivan, 1970). He warned that statements made by clients are not rigidly fixed to meaning. They are, at best, approximations of what is meant.

Additionally, clinicians may give certain verbal or nonverbal cues that inadvertently influence the patient to answer in a particular manner (Rogers, 2001). Clinicians may appear uncomfortable with certain topics of concern, thereby losing valuable information (Rogers, 2001). Further, certain clinicians may be biased toward certain diagnoses (Lanyon & Goodstein, 1982).

Human influences are reduced when an interview is highly structured. Reliability and validity are increased when the error source is lessened. Such reduction in bias has been documented in several studies with highly structured interviews (Bruss, Gruenberg, Goldstein, & Barber, 1994; Mannuzza, Fryer, & Klein, 1993).

Clinical interviews play a large role in psychodiagnostic assessment. Interviews provide an opportunity to gain information directly from the patient regarding the presence and severity of psychopathology. They allow clinicians to evaluate whether or
not the individual is experience emotional difficulties and to determine possible Axis I

The next section will discuss the two main types of interviews, structured and
unstructured, highlighting the benefits of the structured approach. It will define and
discuss the purpose of emotional status examinations as structured interviews, as well as
their diagnostic utility as compared to the MMPI. Finally, the composition of the Dean-
Woodcock Emotional Status Examination, which is a component of the clinical interview
from the Dean-Woodcock Neuropsychological Battery, will be explored.

**Types of clinical interviews.** There are two general types of clinical interviews,
structured and unstructured. The structured interview is a technique for quantitatively
evaluating the degree to which psychopathology is present in a patient (Perloff, Craft, &
Perloff, 1984). It consists of a standardized interview schedule on which a trained
interviewer observes and “scores” the patient as the interview progresses. Structured
interviews define what is asked and how it is asked, thus influencing patient responses
through the choice of a particular line of inquiry (Hodges, 1993).

Unstructured interviews do not dictate the order or wording of questions to be
asked. They allow for more flexibility in the interview process. The interviewer
determines the questions asked, the framing of the questions, and acceptable responses.
It allows the assessment to be tailored to the individual needs of the patient and relies
heavily on the individual clinician’s orientation and expertise (Kamphaus & Frick, 2002).
The content of structured and unstructured interviews varies. In neuropsychology, the
interview focuses on emotional status, general and medical history, and cognitive
functioning. Of interest in this study is the emotional status of the patient.
Strengths of structured interviews over unstructured interviews. Assessment as part of the diagnostic process has been advocated as a method by which practitioners use a scientific approach to their work with patients (Vacc, 1982). The intent of this method is to apply scientific principles, such as drawing from theory and observation to make hypotheses that are then tested during the interview. Clinicians have long embraced the interview as a powerful assessment tool. In the past, numerous interviews have been published for the purposes of research and practical applications (Hodges, 1993). Yet, despite the almost universal endorsement of the approach, the effectiveness and appropriateness of interviews for assessment continued to be questioned due to the unstructured nature of early interviews (Kamphaus & Frick, 2002). Spitzer, Endicott, and Robbins (1978) indicated that unstructured interviews are susceptible to information variance due to the differences in the quantity and quality of information that is obtained from the same patient by different examiners. Edelbrock and Costello (1990) reported that unstructured interviews are one of the least trustworthy assessment procedures. Unstructured interviews often demonstrate lower levels of reliability than other assessment procedures (Hodges, 1993).

Researchers have sought to refine the interview in order to better integrate clinical assessment and diagnosis (Rogers, 2001). The result of this refinement is the structured interview. Structured interviews were developed largely to assist clinicians in the area of differential diagnosis (Segal, Corcoran, Coughlin, 2002). The structured format is particularly helpful in current mental health practice, which stresses diagnostic precision, time-limited intervention, and demonstrable effectiveness (Piotrowski & Keller, 1992). In the cases of assessment in which an initial diagnosis needs to be made, it is advisable
to use formal and well-standardized interviews (Hebben & Milberg, 2002). Standardized interviews can help clinicians organize their clinical judgment, and thereby improve diagnostic precision and create effective interventions to meet patient treatment goals (Rogers, 2001). A structured approach helps the examiner to avoid missing major areas and to gather further information that will establish expectations for level of performance on subsequent tasks (Goldstein, Nussbaum, & Beers, 1998).

Structured interviews are valuable measures (Edelbrock & Costello, 1990). Two meta-analyses have concluded that structured interviews are more valid than unstructured interviews (Wiesner & Cronshaw, 1988; Wright, Lichtenfels, & Pursell, 1989). Saghir (1971) found that diagnostic information gathered using a structured interview tended to be more uniform and comprehensive in its coverage in psychopathology than that gathered in an unstructured interview. Structured interviews provide for the collection of objective data during the diagnostic process, thus enabling direct comparisons to be made among patients or research participants (Edelbrock & Costello, 1990). The diagnoses generated by structured interviews are based on DSM diagnostic criteria, which are used universally by clinicians to make diagnoses (Hodges, 1993).

There are a growing number of standardized, reliable, and valid instruments to assess psychopathology. The utilization of such instruments is becoming the standard for documenting clinical change in psychiatric disorders (Rogers, 2001). The Food and Drug administration (FDA) routinely demands the use of standardized assessment to document the comparative efficacy of old and new psychotropic medications (Othmer & Othmer, 2001). With the increased demand for efficacy of treatment, such developments in
diagnostic assessment appear to be a positive step forward in the field of mental health (Othmer & Othmer, 2001).

**Structured interviews.** As the reliance on the DSM taxonomy system increased, the need developed for more valid and reliable ways to assess patients for the purposes of diagnosis and treatment planning. Before the more recent versions of the Diagnostic and Statistical Manual and the increased interest in measuring the effectiveness of interventions, structured interviews were not commonly used in clinical settings. Instead, unstructured interviews that were dependent upon the clinician, his/her experience and training, and goals were utilized. The content of the information gathered reflected great variability. Currently, however, structured interviews provide clinicians with relevant information that supports clinical diagnoses, influences treatment planning, and summarizes patient-reported progress (Edelbrock & Costello, 1990).

Structured interviews typically have been used for diagnostic or descriptive assessment (Edelbrock & Bostello, 1990). In 1980 the *DSM-III* provided a differentiated categorization of mental health disorders that was explicit in criteria, reshaping the way structured interviews were designed and implemented (Hodges, 1993). The advent of the *DSM-III* led structured interviews to strive to provide a specific DSM diagnosis at the conclusion of the assessment (Hodges, 1993).

Structured interviews that are utilized for descriptive purposes typically indicate the degree of psychopathology present and pertinent information regarding the presenting problem. Although diagnostic interviews are often more standardized, descriptive assessment aids with the differentiation and classification of diagnoses. For example, descriptive interviews may help differentiate between Dysthymia and Major Depression.
They are not based on any particular theoretical framework, relying instead on face validity (Edelbrock & Costello, 1990).

**Types of structured interviews.** Key elements in a structured interview include: clarity, concise introduction, short items, no compound items, constant proportion of items to criterion, no technical jargon, no negatively phrased items eliciting a positive response, and clear instructions/training for administration and scoring (Sudman & Bradburn, 1982). As an assessment instrument, the amount of clinical knowledge required by the examiner varies. Some structured interviews require vast knowledge of the field, whereas others require only basic training on interviewing techniques (Gutterman, O’Brien, & Young, 1987). The structured interview usually requires the examiner to ask the question in an ordered sequence and involves a series of interactions between examiner and patient, usually in a question-and-answer format, that is used as the basis for making inferences. The variability in approach defines the instrument as either a highly structured interview or a semi-structured interview (Edelbrock & Costello, 1990).

Highly structured interviews specify the exact order and wording of questions and provide explicit rules for rating and recording a patient’s responses. The intent of the highly structured interview is that all clinicians ask exactly the same questions in the same order and evaluate and record the responses in the same way. The semi-structured interview is less restrictive and permits the clinician flexibility in the process while still guiding the interview.

**Reliability and validity of structured interviews.** Reliability and validity are important psychometric qualities to consider when critically evaluating diagnoses.
generated by structured interviews (Segal & Coolidge, 2003). The reliability of structured interviews is generally based on research of inter-rater reliability (Farmer & Chapman, 2002). Researchers have examined specific structured interview instruments to determine their inter-rater reliability (e.g., Diagnostic Interview for Social & Communication Disorders, Geriatric Depression Scale, Structured Clinical Interview for DSM, Computer Assisted Diagnostic Interview, Structured Expert Interview for Anorexic and Bulimic Syndromes). The inter-rater reliability for each instrument was found to be high (.62-1.00; Fichter & Quadflieg, 2001; Gutterman, O’Brien, & Young, 1987; Miller, 2001; Wing, Leekam, Libby, Gould, & Larcombe, 2002; Wong, et al., 2002).

Specifically, Wong, et al., (2002) compared the inter-rater reliability between a standardized and a non-standardized administration of the Geriatric Depression Scale, indicating that the standardized procedure in administration and scoring improved the inter-rater reliability by structuring the wording and order of questions asked. Miller (2001) compared the results of the structured Computer Assisted Diagnostic Interview (CADI) administered within the emergency room and then in the inpatient unit to results of the unstructured Traditional Diagnostic Assessment (TDA). The structured CADI revealed an inter-rater agreement of 79.5% as compared to the unstructured inter-rater reliability of the TDA (45.5%).

Test-retest reliability is another method employed to assess the reliability of interviews. Paget (1984) reported that test-retest reliability was often missing in studies of structured interviews. The lack of test-retest reliability leaves unanswered the issue of instruments providing consistent results over time (Bucholz, et al., 1994).
After reviewing three structured instruments, Gutterman, et al., (1987) reported that they had good test-retest reliability, typically ranging from .63 to 1.00, but only in relation to certain diagnoses. Hodges’ (1993) examination of similar instruments demonstrated varied test-retest reliability, with correlations ranging from .39 to .89 and shifting significantly depending on the specific diagnosis being studied. More research to determine the test-retest reliability of structured interviews is warranted (Hodges, 1993).

The most prevalent data on the psychometric properties of structured interviews concerns concurrent validity (Blake, et al., 1995; Cyranowski, et al., 2002; Freedland, et al., 2002; Gunderson, Phillips, Tiebwasser, & Hirschfield, 1994). Concurrent validity has been demonstrated for several structured interviews (e.g., Beck Depression Inventory, Depression Interview and Structured Hamilton, SCID, SCISAS). For example, the Depression Interview and Structured Hamilton (DISH) is a semi-structured interview developed by Freedland and colleagues (2002). Good test-retest reliability (93%) was found for the interview between clinicians and nurses. When the DISH was compared to the Structured Clinical Interview for DSM-IV (SCID) in a group of 57 inpatients, an agreement in diagnoses was obtained by the two measures in 88% of the cases. The disagreements between the DISH and the SCID were, in every instance, between adjacent categories (i.e. major vs. minor depression, or minor depression vs. no depression).

Some of the best reported psychometric data for diagnostic validity studies were for clinician-generated diagnostic interviews and diagnoses at time of discharge, based on all information. Hodges (1993) reported moderate correlations, depending on the diagnostic category.
Problems exist in establishing reliability and validity of interview instruments. Much of the research reported about structured interviews has been conducted by the developers of the instruments. This could potentially lead to bias in the study and in the interpretation of results. Also, an examination of studies concerning structured interviews indicates that little work has been done comparing different structured interviews with the same population. In order for structured interviews to be viewed as viable assessment instruments by the psychometric community, researchers need to focus on using larger sample sizes and determining one or more established instruments for comparison studies (e.g., Harner & Heal, 1993; McCanse, 1995). Also, analyzing structured interviews as a whole rather than analyzing each interview as a separate entity should be explored further through meta-analyses. This may help to promote the view of structured interviews as sound assessment tools.

There are limitations in the study of validity in particular. Frequently the interview schedules include subsections that are based on the assumption that the instruments are measuring a variety of symptoms or constructs, though this has not been validated (Paget, 1984). In addition, most studies used small sample sizes that are restrictive in representation (Harner & Heal, 1993; McCanse, 1995; Stone & Hogan, 1993). Accordingly, generalizability of conclusions may be limited. This may be due in part to many structured clinical interviews being labor and time intensive. Therefore, there is a strong need for a reliable and valid structured interview that is easy to administer and does not require a lot of time.
Emotional Status Examinations

An integral part of a structured clinical interview is a measure of emotional adjustment. The emotional status examination measures emotional adjustment and is as crucial to the field of mental health as the physical examination is to other areas of medicine (Robinson, 2000). It explores the patient’s subjective perception of her or her feelings. The clinician then must consider the patient’s presenting complaints as well as a particular set of symptoms that validate specific a priori assumptions about human behavior and psychopathology (Schwartz, 1989). Goldstein, Nussbaum, and Beers (1998) stated that “the information gathered should allow the examiner to tentatively formulate (1) some expectations about the general level of . . . functioning of the patient, (2) the syndromatic presentation of the patient, and (3) etiological . . . differential diagnosis,” (p. 21).

The content of emotional status examinations is generally based on the criteria specified in the DSM (Hodges, 1993). The questions asked are directly linked to the diagnostic criteria of the various emotional disorders. In the interview format of emotional status examinations, a stem question leads to set of questioning. The detail and structure of the questions depend on the degree of standardization of the interview. If the stem question is answered affirmatively, then follow-up questions are asked to determine other relevant parameters such as frequency, severity, duration, and level of impairment. This allows the clinician to gather pertinent information necessary to properly diagnose the patient. Other types of emotional status examinations follow a self-report format that allows the patient to rate behaviors associated with their emotional functioning (e.g.,
Zung Depression Rating Scale). This format is not as comprehensive, as follow-up questions are not employed to help determine issues related to diagnosis.

The emotional status examination as a structured interview is of primary interest in this study. Structured emotional status examinations incorporate all of the important components of structured interviews discussed earlier (e.g., straightforward wording enabling clinicians to gather pertinent diagnostic information). Beyond gathering diagnostic information in a systematic way, the patient’s current mental and emotional state, including thinking, affect, and behavior, are recorded (Schwartz, 1989).

The primary advantage of structured emotional status examinations is their usefulness in obtaining important parameters of a patient’s behavior that is not typically assessed in most behavior rating scales. Specifically, most provide questions that elicit information on the duration of specific symptoms and the age of onset. They also allow the clinician to determine temporal sequencing among behaviors and feelings. For example, it is important in the assessment of depression to determine whether periods of sadness occurred contiguously with other behaviors associated with depression, such as sleep disturbances, eating disorders, or thoughts of death (Kazdin, 1988). Another important parameter of structured emotional status examinations is the level of impairment associated with the behaviors and feelings being reported. Questions regarding the degree to which a patient’s difficulties are affecting his or her functioning in major life areas (e.g., at home, work, and with peers) are crucial to DSM diagnosis.

In addition to assessing important parameters of behavior and feelings, structured emotional status examinations enhance the correspondence between assessment techniques and diagnostic criteria (Kamphaus & Frick, 2002). The tie between
assessment and the DSM diagnostic criteria is advantageous for several reasons. First, it promotes revisions of the emotional status examinations to correspond with advances in knowledge of the basic characteristics of emotional disorders. Second, it allows clinicians to make diagnoses based on strict adherence to diagnostic criteria. Clinicians utilize the information gathered from the assessment to make diagnoses for theoretical, empirical or practical reasons. Too often diagnoses are made based on information that does not directly assess the diagnostic criteria (Kamphaus & Frick, 2002). As a result, the meaning of the diagnosis is ambiguous. It is important to note, however, that using the DSM criteria to determine diagnosis limits the emotional status examination in the same ways the DSM itself is limited. Since some diagnostic categories have been poorly validated, this dependency is not always beneficial to the assessment process (Kamphaus & Frick, 2002). Still, the use of the DSM criteria to determine areas of assessment provides the most meaningful, generalizable diagnostic information.

Of interest in this study is the reported emotional status of patients and how this data correlates with the MMPI. Determining increased validity for emotional status examinations is useful in emergent situations when the MMPI is invalid, with resistant patients, or when the patient is unable to complete the MMPI. A trained clinician’s ability to observe and recognize known indicators of pathological entities, the structured format, as well as the relative independence of the observations from the verbalized content of the interview, allows for the validity and reliability of this process as a diagnostic tool (Schwartz, 1989).

**Diagnostic utility of structured interviews/emotional status examinations.** An approach commonly used to measure the validity of an instrument is to compare
diagnoses generated by a structured interview with diagnoses from another source (Segal & Coolidge, 2003). Research exploring the predictive ability of emotional status examinations primarily focuses on overall emotional status, or depression and anxiety in particular (e.g., Braaten, et al., 2001; Sunday, et al., 2001). The instrument assessing emotional status is often compared to a self-report personality inventory. There is great variety in the personality inventories used for comparison (e.g., Geriatric Depression Scale, Millon Clinical Multiaxial Inventory, Personality Diagnostic Questionnaire, Symptom Check List), as well as the emotional status examinations themselves (e.g., Apparent Emotional Rating Scale, Hamilton Rating Scales for Anxiety and Depression, International Personality Disorder Examination, Structured Clinical Interview for DSM, Structured Interview for DSM-IV Personality).

Research looking at the diagnostic ability of structured interviews (e.g., Structured Clinical Interview for DSM) has found them to be useful in detecting Axis I affective, anxiety, and depressive disorders (e.g., Braaten, et al., 2001; Sunday, et al., 2001). Braaten and colleagues (2001) utilized structured diagnostic interviews to compare the reported symptoms of depression between teens and their parents. The findings suggested that the youth self-reports tended to identify a more mild form of depression than that of parents’ reported observations. Sunday and colleagues (2001) compared two versions of the same interview used with a group of female inpatients. They found complete agreement between the SCID for the DSM-III-R and the DSM-IV for all Axis I disorders except anorexia nervosa (79%), alcohol abuse/dependence (80%), and substance abuse/dependence (97-99%). The agreement between versions for Axis II diagnoses varied by personality disorder (92-100%). Lower levels of agreement for
personality disorders across measures is believed to be due to the nature of personality disorders rather than a fault with the method of assessment (Farmer & Chapman, 2002).

In looking at overall emotional status, structured interviews were frequently compared to the Millon Clinical Multiaxial Inventory (MCMI; e.g., Hunt & Andrews, 1993; Messina, 2000; Messina, Wish, Hoffman, & Nemes, 2001; Miller, Streiner, & Parkinson, 1992; Soldz, Budman, Demby, & Merry, 1993). Low levels of agreement were found between the structured interviews and the MCMI, which may be due in part to the different types of information collected by each. The Structured Clinical Interview for DSM (SCID), for example, emphasized observable criteria, while the Millon Clinical Multiaxial Inventory emphasized pathological traits. Segal (1997) suggests that the poor concordance rates are a limitation of the self-report inventory rather than of the structured interview because the former tends to result in substantial false positives.

Sunday and colleagues (2001) examined the differences in Axis I and Axis II diagnoses using the SCID for the DSM-III-R and the SCID for the DSM-IV. The concordance between the two interviews was excellent for all axis I disorders except anorexia nervosa and substance abuse. The agreement on Axis II varied by personality disorder. The discrepancies were attributed to the change in diagnostic criteria between the two versions of the DSM rather than a fault with the interviews.

In assessing depression, the SCID has demonstrated good diagnostic concordance with self-report measures (Foley, Neale, & Kendler, 2001). Specifically, Manne and colleagues (2001) compared the SCID-I to the Beck Anxiety Inventory (BAI and the Beck Depression Inventory (BDI), both self-report measures. They evaluated anxiety and depression in 115 mothers of children undergoing bone marrow transplants. The
SCID demonstrated more predictive accuracy in assessing Generalized Anxiety Disorder and Panic Disorder as compared to the BAI; whereas, the SCID and the BDI were comparable in their ability to assess Major Depressive Disorder.

When an interview and self-report measure that were designed together were utilized, a high level of agreement was found (Dell’Osso et al., 2002). The level of agreement between the interview and the self-report formats of the Structured Clinical Interview for Social Anxiety Spectrum (SCI-SHY) and the Structured Clinical Interview for Obsessive Compulsive Spectrum (SCI-OBS) was analyzed in a group of 30 psychiatric patients and 10 control subjects. Agreement between the two versions was very high, indicating high predictive validity.

As mentioned previously, there has been less success with the diagnosis of personality disorders. This is thought to be attributed to problems with the assessment and/or conceptualization of some personality disorder categories, rather than the method used to assess the personality disorder criteria (Farmer & Chapman, 2002). Farmer and Chapman (2002) compared the Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II) with the Structured Clinical Interview for DSM-IV Axis II Personality Disorders Questionnaire (SCID-II-PQ), a self-report measure. Excellent inter-rater reliability was found on the SCID-II. However, the SCID-II-PQ produced a high rate of false-positive personality disorder diagnoses relative to the SCID-II findings.

Similar results were found when evaluating general distress (Heyman, Feldbau-Kohn, Ehrensaft, Langhinrichsen-Rohling, & O’Leary, 2001). The Dyadic Adjustment Scale (DAS, self-report measure) was compared to an author-constructed face valid structure diagnostic interview to assess the perceived potential for physical abuse in a
relationship. The structured interview had modest agreement with the DAS, but was able to more accurately diagnose distress compared to the DAS.

**Diagnostic utility as compared to the MMPI.** Of particular interest to this study is how structured emotional status examinations compare to MMPI profiles. In general, findings suggest that both forms of assessment demonstrated a statistically reliable degree of association with the DSM-based categories of psychopathology and with each other (Griffiths & Gillingham, 1978; Huckaby, Kohler, Garner, & Steiner, 1998; Kilpatrick, 1978; Liskow, Penick, Powell, & Haefele, 1986; Lucas, 1974; Newmark, Gentry, Simpson, & Jones, 1978; Oberstone & Sukoneck, 1976; Stangl, Pfohl, Zimmerman, Bowers, & Corentahl, 1985; Svanum & McGrew, 1996). This suggests high predictive validity.

Svanum and McGrew (1996) compared the MMPI and MMPI-2 profiles of 690 undergraduate college students to the DSM-III-R diagnoses derived from the DSM-III-R diagnoses derived from the Diagnostic Interview Schedule (DIS). Procedures of multiple regression/correlational analysis were used to test and compare the predictive ability of the MMPI and MMPI-2. Both versions of the MMPI demonstrated a statistically reliable degree of association with the diagnoses and overall emotional adjustment obtained from the DIS.

Similar results were obtained by Stangl and colleagues (1985). They utilized the Structured Interview for the DSM-III Personality Disorders (SIDP) and the MMPI with 63 inpatients. The results showed that patients with personality disorders according to the SIDP scored higher than those without personality disorders on all MMPI scales, indicating good correlation between the SIDP and the MMPI.
The MMPI and the Mental Status Schedule, a structured interview, were used to assess the existing diagnoses of schizophrenia in 469 inpatients (Newmark, Gentry, Simpson, & Jones, 1978). Results found that 72% of those reliably diagnosed as schizophrenic with the interview were detected on the MMPI. Only 5.5% of non-schizophrenic patients obtained a similar MMPI profile. Newmark, Gentry, and Whitt (1983) attempted to replicate these findings in an adolescent population (89 inpatients). However, they were not successful; only 23% of the sample obtained an MMPI profile of schizophrenia. This was attributed to the different symptoms of schizophrenia in adolescents.

Specific scales on the MMPI have been found to have good concordance with measures of emotional status. The Weinberger Adjustment Inventory (WAI) was employed with 178 male, incarcerated juvenile offenders (Huckaby, Kohler, Garner, & Steiner, 1998). The distress dimension (including consideration of others, suppression of aggression, impulse control, and responsibility) of the WAI was significantly and positively associated with scale 2 (symptomatic depression, .59), 4 (social maladjustment, .52) and 7 (obsessive-compulsive behavior, .56) of the MMPI. The restraint dimension (including anxiety, depression, self-esteem, and low well-being) was significantly and positively associated with the Lie (.37) and K (.32) scales of the MMPI. The results suggest that the WAI is a useful measure of psychopathology.

Kilpatrick (1978) found similar levels of reported depression and anxiety on the MMPI and emotional status examinations. He utilized the Profile of Mood States Scale (POMS) and the Speilberger State-Trait Anxiety Inventory (STAI) with a group of 154 males hospitalized for alcohol detoxification. The subjects reported themselves to be
high in anxiety and depression across measures. Johnson, Klinger, and Williams (1977) utilized the MMPI and the Current and Past Psychopathology Scale to determine if the MMPI validity indicators were equally as able to assess negative behaviors as structured interviews. A group of 228 veterans receiving treatment from a VA hospital participated in the study. A statistically significant relationship between the F scale of the MMPI and ratings of “fake bad” interview behavior were found.

Studies have also been conducted using the MMPI and various structured interviews/emotional status examinations to determine differences between groups of subjects. They have been shown to be useful in this manner (Liskow, Penick, Powell, & Haefele, 1986; Lucas, 1974; Oberstone & Sukoneck, 1976). For example, Liskow, Penick, Powell, and Haefele (1986) compared a group of psychiatric inpatients identified as having Briquet’s syndrome with the Psychiatric Diagnostic Interview (PDI) to a similar control group. A comparison was made between the two groups on their PDI and MMPI scores. Briquet’s patients averaged more than twice as many positive syndromes as the controls (4.7 vs. 2.3). On the MMPI, they averaged 6.3 of 13 scales elevated about 70 compared with 3.0 for the control group, a significant difference. While no correlation was found between the number of PDI syndromes positive and the number of MMPI scale scores higher than 70, the Briquet’s syndrome patients did report higher levels of pathology on both measures than did the control group. Psychological adjustment and lifestyle of single heterosexual and homosexual females was evaluated by comparing structured interview and MMPI results of the two groups (Oberstone & Sukoneck, 1976). No major differences were found between groups.
Research with unstructured interviews as correlated with the MMPI, though, has shown less reliability (Stangl, Pfohl, Zimmerman, Bowers, & Corentahl, 1985). Because of the abundant discrepancies in the research involving structured versus unstructured interviews and their correlation with the MMPI, this study utilizes a structured interview, the Dean-Woodcock Emotional Status Examination.

**Dean-Woodcock Emotional Status Examination**

One specific structured emotional status examination that is based on DSM diagnostic criteria is the Dean Woodcock Emotional Status Examination (D-WESE). The D-WESE was constructed to be incorporated into the Dean-Woodcock Neuropsychology Battery as a structured interview component that is to be utilized with each administration. The Dean-Woodcock Structured Interview includes the D-WESE that assesses emotional functioning. Of interest to this study is the strength of the correlation between the D-WESE and the well-established MMPI.

**Construction of the Dean-Woodcock Emotional Status Examination.** The D-WESE was constructed with the pertinent components of interviews and mental status examinations in mind (Dean & Woodcock, 2003). The key elements of structured interviews were included (e.g., clarity, short items) in order to increase reliability of the measure. Once the interview was developed, peer evaluations of the experimental protocol were obtained. It was then field-tested on a modest sample (Dean & Woodcock, 2003).

The D-WESE assesses the presence and severity of psychiatric symptoms in a highly structured manner. These symptoms include depression, mania, attention, obsessive/compulsive behavior, anxiety, somatization, aggression, and paranoia. The
development of the D-WESE followed the scientific approach recommended by Vacc (1982), attending to the taxonomy of the DDSM in order to make predictions about diagnoses. The questions are based on the discrimination criteria of the DSM in order to aid in differential diagnosis, which is thought to improve the validity of the measure (Dean & Woodcock, 2003). In the administration of the D-WESE, the wording and order of the questions are dictated. Patients are asked to admit or deny the presence of each symptom. For example, “Do you feel depressed, sad, blue, or ‘down in the dumps’?” If the patient reports experiencing the symptom, he/she is then asked to rate the severity of the experience as mild, moderate, or severe. The question is then followed with other questions pertaining to specific symptoms associated with depression (e.g., suicidal thoughts, sleep disturbances) and acknowledged symptoms are recorded and rated.

Clinical observations either support or refute information provided by the patient and are documented. These include: (a) orientation, (b) attention, (c) comprehension, (d) speech, (e) consciousness, (f) grooming, (g) appearance, (h) perceptual disturbance, (i) affect, (k) agitation, (l) flow of thought, (m) mental content, (n) task persistence, (o) interaction, (p) insight, and (q) judgment. The clinical observations section is semi-structured in how it is documented. Options are provided that relate to the area of interest and limit the acceptable responses, but more than one option can often be indicated. For example, attention can be recorded as focused, needed redirection, distracted easily, or appeared impaired. A patient may have needed redirection and also may have been distracted easily, in which case both options would be designated.

In this study, each of the D-WESE items was scored on a 5-point Likert-type scale to represent presence or denial and severity of symptoms (i.e., denied=1, yes=2,
infrequently=3, frequently=4, very frequently=5). This procedure allowed for a combination of the yes/no and frequency responses to the D-WESE so that one variable could be used to represent each item.

**Rationale for selection of the D-WESE.** The D-WESE possesses many strengths in its attempt to assess emotional functioning. In question is whether the emotional status examination can be made sufficiently reliable and valid to be viewed as a scientifically acceptable assessment and/or research instrument (Segal & Coolidge, 2003). The answer, in part, involves structuring the interview process and limiting the variability in the question/response transactions between the examiner and the patient. Inherent in this process is defining the content to be assessed, limiting the order and wording of the questions, and standardizing how responses are evaluated, recorded, and interpreted (Sudman & Bradburn, 1982).

These necessary components were included in the D-WESE’s construction (e.g., clear, straightforward questions). The D-WESE is specifically designed to correspond to the *DSM-IV* diagnostic criteria. Therefore, the meaning of the diagnoses derived from the D-WESE is universal across clinicians. In addition to diagnostic information, the D-WESE elicits data regarding the intensity of specific emotions and the level of impairment associated with the behaviors and feelings, which provides descriptive information that is pertinent to diagnoses, treatment, and recovery. Clinical observations are recorded and used to either support or refute the patient’s self-report information, further aiding in diagnosis and treatment. Few emotional status examinations were found to include sections designed to document observations in such a semi-structured manner.
The nature of the D-WESE and ease of documentation on the protocol allow a wide breadth of information to be covered in a short period of time. In a time when diagnostic precision and time-limited intervention is necessary, the D-WESE is an efficient assessment tool that can be completed in a timely manner (Dean & Woodcock, 2003). While the highly structured nature of the assessment improves its reliability, it must be noted that it also limits the flexibility that the clinician has in administration. It restricts the order and wording of questions limiting further exploration of additional information. Although this may be considered a criticism of the D-WESE, it also improves the reliability and validity of the information gathered. Another concern regarding the use of the D-WESE is that it is a new measure with little research coverage. The D-WESE needs to be evaluated to identify its utility in neuropsychological assessment. Preliminary research demonstrates significant levels of shared variability between the D-WESE and the MMPI (Galloway-Sharp, 2004). However, further research is necessary to determine and analyze the amount of shared and unshared variance between the D-WESE and the MMPI.

Summary

Two fundamental components of emotional assessment are objective personality measures such as the MMPI and clinical interviews. Despite the criticisms of the MMPI, it has been and continues to be widely used in clinical practice as well as research. Further, it has been used to validate a number of diagnostic interviews and emotional status examinations.

Clinical interviews involve face-to-face dialogue between patient and clinician used to assess the presence and severity of psychopathology. These interviews can be
structured or unstructured in their format. Emotional status examination is a major component of clinical interviewing. Providing structure to clinical interviews significantly improves their reliability and validity. Structured emotional status examinations provide clinicians with relevant information that supports clinical diagnoses, influences treatment planning, and summarizes patient-reported progress.

Structured clinical interviews and emotional status examinations generate diagnoses based on DSM diagnostic criteria and have been found to be useful in detecting Axis I affective, anxiety, and depressive disorders. They have also demonstrated a statistically reliable degree of association with the MMPI.

The Dean-Woodcock Neuropsychology Battery (D-WNB) incorporates a structured interview called the Dean-Woodcock Structured Interview (D-WSI), which includes a structured emotional status examination, the Dean-Woodcock Emotional Status Examination (D-WESE). The D-WESE assesses the presence and severity of psychiatric symptoms in a structured manner, in order to determine the presence of DSM Axis I diagnose. For the purposes of this study, the D-WESE is compared to the MMPI in order to determine diagnostic utility. Specifically, the amount of shared and unshared variability between the D-WESE and the MMPI was measured and analyzed.
CHAPTER III

METHOD

Participants
The participants \((n = 207)\) included 82 male and 125 female White patients who had been referred for a neuropsychological evaluation and treatment at a large outpatient Midwestern neurology practice. The age of participants ranged from 13 to 96 \((M = 56.53, SD = 21.04)\). The participants in this study varied in level of impairment from mild or no impairment to marked neuropsychological impairment.

All participants were administered a standard battery of measures by a licensed psychologist or technician or student supervised by a licensed psychologist. An emotional status exam, assessments of neuropsychological functioning, social-emotional functioning, and cognitive ability were administered during each evaluation. In order to maintain patient confidentiality, identifying information of patients was replaced with identification numbers.

Instrumentation

Dean-Woodcock Emotional Status Examination. The Dean-Woodcock Emotional Status Examination (D-WESE) is a well-constructed, individual component of the Dean-Woodcock Neuropsychological Battery. The D-WESE seeks the presence and
severity of psychiatric symptoms in a structured manner to determine the presence of diagnoses as classified on Axis I of the *Diagnostic and Statistical Manual – Fourth Edition – Text Revision* (DSM-IV-TR). The *D-WESE* seeks important information regarding the symptoms associated with behaviors and feelings and provides descriptive information necessary for diagnosis, treatment, and recovery. It was originally written in 1994 and later revised in 1998 (Dean, 1998) to include questions about executive functioning. Because the *D-WESE* is a newly developed measure, more information requiring its psychometric properties is needed. Initial research revealed eleven factors: intellectual/executive functioning, anxiety symptoms, asocial behaviors, depressive features, somatic symptoms, inattention, unstable mood/psychotic, post-traumatic symptoms, energy level, appetite/sleep, and rebelliousness (Stage, in press).

The *D-WESE* is administered in the form of a structured interview. The directions for administration are included in the interview form. There are 50 items that can be asked directly to the patient or indirectly through an informant. Each question represents a symptom of one or more psychiatric conditions. The patient admits or denies having the symptom. The examiner may then choose to rate the severity of the symptom on a Likert-type scale (e.g., mild, 1-4; moderate, 5-7; severe 8-10) depending upon its perceived importance relative to the referral question.

An exploratory factor analysis of the *D-WESE* was performed, with eleven factors emerging (Galloway-Sharp, 2004, Stage, in press). A Pearson r correlation was conducted using the eleven identified factors of the *D-WESE* and the ten clinical scales of the MMPI (Galloway-Sharp, 2004). Results indicated each *D-WESE* factor was found to be significantly correlated with five to nine of the MMPI Clinical scale, indicating the
factors and scales measured similar information. The Dean-Woodcock Sensory Motor Battery, another component of the Dean-Woodcock Neuropsychological Battery, has been found to have adequate to excellent inter-rater agreement (Woodward, Ridenour, Dean, & Woodcock, 2002). However, little information regarding specific reliability of the D-WESE is available.

**Minnesota Multiphasic Personality Inventory.** The *Minnesota Multiphasic Personality Inventory* (MMPI) was developed by Hathaway and McKinley (1940) as a paper-and-pencil personality inventory to provide an efficient way of arriving at a psychodiagnosis (Graham, 1999). The MMPI consists of 566 self-reference statements. For each item, the respondent indicates whether the statement is true or false as applied to him/herself. Responses are scored according to a set of empirical keys, which yield scores on ten clinical scales and three validity scales. The clinical scales measured by the MMPI include Hypochondriasis, Depression, Conversion Hysteira, Psychopathic Deviate, Masculinity-Femininity, Paranoia, Psychasthenia, Schizophrenia, Hypomania, and Social Introversion. The validity scales, which include Lie, Infrequency, and Defensiveness scales, determine test-taker attitude regarding potential over- or underreporting of symptoms. MMPI scores are calculated by summing the number of subject endorsed items that load on the particular scale and converting the raw score into a *T*-score. Clinical and validity scales have an average *T*-score of 50 with a standard deviation of 10. *T*-scores of 65 or above are considered elevated, indicating the patient has endorsed a significant number of items related to the scale. A multitude of research exists regarding validity, reliability, and possible uses of items and scales of the MMPI.
(Hathaway & McKinley, 1942). See Chapter 2 for further information regarding validity and reliability of the MMPI.

**Procedure**

A randomized research design was used to select participants from a group of patients who were referred for neuropsychological assessment as part of a neurological evaluation. Participants completed consent forms at the time of the evaluation to allow their results to be utilized in research. Participants were individually administered the *Dean-Woodcock Emotional Status Examination* (D-WESE) as part of the *Dean-Woodcock Neuropsychological Battery* (D-WNB). The participants were then administered the *Minnesota Multiphasic Personality Inventory* (MMPI) in keeping with standardization.

**Data Analysis**

The examination and inventory scores were compiled and descriptive and inferential statistics were calculated. The scored items of the D-WESE were grouped into eleven factors derived by a confirmatory factor analysis by Stage (in press) and summed to provide a factor score. Ranges for each of the factor scores are provided below:

- Intellectual/Executive Functioning: 0-8
- Anxiety Symptoms: 0-7
- Asocial Behaviors: 0-7
- Depressive Features: 0-5
- Somatic Symptoms: 0-3
- Inattention: 0-4
- Unstable Mood/Psychotic Symptoms: 0-4
Next, a canonical correlation was performed to determine the levels of shared and unshared variability between the eleven factors of the D-WESE and the ten clinical scales of the MMPI. Canonical correlation is a statistical technique that enables the assessment of the degree of linear relationship between two sets of variables. Canonical correlation analysis represents the highest level of the general linear model (Sherry & Henson, 2005). The technique relies on the extraction of linear combinations within each set of variables in a manner that allows for maximizing the correlation between the two sets (Lee, 1978).

The results of a canonical correlation of these data provided information regarding the relationship between each of the factors/scales and the strength of each factor in relation to the overall correlation. Based on the results of previous research, (Galloway-Sharp, 2004; Stage, in press), it was hypothesized that the MMPI and the D-WESE would be found to share a significant level of variance, suggesting support for the use of the D-WESE in a clinical setting. Although this mutual shared variability will likely exist, past research also suggests unique variability between measures (Galloway-Sharp, 2004). The MMPI and the D-WESE were both designed to measure emotional symptoms, indicating they will correlate highly as measures of overall distress (Galloway-Sharp, 2004). However, the MMPI was originally developed to assess the degree to which patient disorders were organic (biological) or functional (situational)
(Van Dorsten & Weisberg, 2011). Therefore, the data provided by the clinical scales and codetypes, which are the patient’s two most elevated clinical scales, provide an overall picture of the patient’s profile and related behaviors. In contrast, the D-WESE measures specific symptoms and symptom groups from the *DSM-IV-TR*, which is necessary for diagnosis and treatment. Because the MMPI provides an overall impression whereas the D-WESE measures specific criteria, it is hypothesized that the D-WESE will account for a significant amount of the variability of the MMPI as well as variance that is not measured by the MMPI. This would indicate that the D-WESE measures some of the same variability as the MMPI, whereas the D-WESE also provides valuable patient information useful in treatment and a better understanding of the MMPI.
A canonical correlation analysis was conducted between ten clinical scales of the MMPI and eleven factor scores of the D-WESE using SPSS CANCORR to estimate the multivariate shared relationship between the two variable sets. Assumptions of linearity of relationship between variables as well as the normality of their distributions were met (Tabachnick & Fidell, 2001). The means and standard deviations for the MMPI and D-WESE appear in Table 1.

Table 1.

Means and Standard Deviations of MMPI Clinical Scales and D-WESE Factor Scores

<table>
<thead>
<tr>
<th>MMPI</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale 1: Hypochondriasis</td>
<td>70.55</td>
<td>14.89</td>
</tr>
<tr>
<td>Scale 2: Depression</td>
<td>72.31</td>
<td>14.07</td>
</tr>
<tr>
<td>Scale 3: Hysteria</td>
<td>68.34</td>
<td>11.99</td>
</tr>
<tr>
<td>Scale 4: Psychopathic Deviate</td>
<td>63.69</td>
<td>12.71</td>
</tr>
<tr>
<td>Scale 5: Masculinity-Femininity</td>
<td>55.69</td>
<td>11.02</td>
</tr>
<tr>
<td>MMPI</td>
<td>Mean</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Scale 6: Paranoia</td>
<td>61.94</td>
<td>13.18</td>
</tr>
<tr>
<td>Scale 7: Psychasthenia</td>
<td>67.43</td>
<td>13.75</td>
</tr>
<tr>
<td>Scale 8: Schizophrenia</td>
<td>70.63</td>
<td>16.75</td>
</tr>
<tr>
<td>Scale 9: Hypomania</td>
<td>56.20</td>
<td>12.58</td>
</tr>
<tr>
<td>Scale 0: Social Introversion</td>
<td>61.20</td>
<td>8.51</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D-WESE</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive Functioning</td>
<td>15.74</td>
<td>6.46</td>
</tr>
<tr>
<td>Anxiety Symptoms</td>
<td>13.62</td>
<td>5.70</td>
</tr>
<tr>
<td>Asocial Behaviors</td>
<td>9.56</td>
<td>3.82</td>
</tr>
<tr>
<td>Depressive Features</td>
<td>10.72</td>
<td>4.17</td>
</tr>
<tr>
<td>Somatic Symptoms</td>
<td>6.85</td>
<td>3.48</td>
</tr>
<tr>
<td>Inattention</td>
<td>10.27</td>
<td>4.49</td>
</tr>
<tr>
<td>Psychotic Symptoms</td>
<td>5.21</td>
<td>2.16</td>
</tr>
<tr>
<td>Posttraumatic Symptoms</td>
<td>4.29</td>
<td>1.75</td>
</tr>
<tr>
<td>Energy Level</td>
<td>4.01</td>
<td>1.28</td>
</tr>
<tr>
<td>Appetite/Sleep</td>
<td>4.60</td>
<td>1.75</td>
</tr>
<tr>
<td>Rebelliousness</td>
<td>2.42</td>
<td>1.02</td>
</tr>
</tbody>
</table>

\( n = 207 \)
The canonical correlation analysis yielded six statistically significant canonical correlations ($\alpha = 0.05$) between the clinical scales of the MMPI and the factor scores of the D-WESE for the neuropsychological patients in the sample. The values of corresponding squared canonical correlations, which represent the proportion of variance shared by the two sets of variables, are listed in Table 2. Three of the significant canonical correlations suggest a strong relationship between the MMPI scales and the D-WESE scores and three suggest a moderate relationship between the sets of variables (Cohen, 1992). It is important to note that though the fourth, fifth, and sixth canonical functions were statistically significant, the percent of variance they explain is relatively low and provides less meaningful information than the first three functions. To examine the nature of the relationship between the MMPI scales and D-WESE factor scores, structure coefficients linking each observed measure with its canonical variable were analyzed. The results for each canonical function appear in Tables 3 and 4.

Table 2.  
*Squared Canonical Correlations between MMPI Clinical Scales and D-WESE Factor Scores*

<table>
<thead>
<tr>
<th>Function</th>
<th>Squared Canonical Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.728</td>
</tr>
<tr>
<td>2</td>
<td>.650</td>
</tr>
<tr>
<td>3</td>
<td>.570</td>
</tr>
<tr>
<td>4</td>
<td>.267</td>
</tr>
<tr>
<td>5</td>
<td>.227</td>
</tr>
</tbody>
</table>
### Table 2 – Continued

<table>
<thead>
<tr>
<th>Function</th>
<th>Squared Canonical Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>.142</td>
</tr>
</tbody>
</table>

### Table 3.
*Structure Coefficients for MMPI Clinical Scales with D-WESE Factor Scores*

<table>
<thead>
<tr>
<th>MMPI Clinical Scale</th>
<th>Function 1</th>
<th>Function 2</th>
<th>Function 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale 1: Hypochondriasis</td>
<td>-.389</td>
<td>.103</td>
<td>.874</td>
</tr>
<tr>
<td>Scale 2: Depression</td>
<td>-.772</td>
<td>-.443</td>
<td>.340</td>
</tr>
<tr>
<td>Scale 3: Hysteria</td>
<td>-.313</td>
<td>-.002</td>
<td>.867</td>
</tr>
<tr>
<td>Scale 4: Psychopathic Deviate</td>
<td>-.680</td>
<td>.203</td>
<td>.205</td>
</tr>
<tr>
<td>Scale 5: Masculinity-Femininity</td>
<td>-.160</td>
<td>.089</td>
<td>-.117</td>
</tr>
<tr>
<td>Scale 6: Paranoia</td>
<td>-.768</td>
<td>.214</td>
<td>-.034</td>
</tr>
<tr>
<td>Scale 7: Psychasthenia</td>
<td>-.884</td>
<td>.013</td>
<td>.172</td>
</tr>
<tr>
<td>Scale 8: Schizophrenia</td>
<td>-.892</td>
<td>.290</td>
<td>.123</td>
</tr>
<tr>
<td>Scale 9: Hypomania</td>
<td>-.505</td>
<td>.591</td>
<td>-.058</td>
</tr>
<tr>
<td>Scale 0: Social Introversion</td>
<td>-.444</td>
<td>-.594</td>
<td>.037</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MMPI Clinical Scale</th>
<th>Function 4</th>
<th>Function 5</th>
<th>Function 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scale 1: Hypochondriasis</td>
<td>.167</td>
<td>.097</td>
<td>-.033</td>
</tr>
<tr>
<td>MMPI Clinical Scale</td>
<td>Function 4</td>
<td>Function 5</td>
<td>Function 6</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>Scale 2: Depression</td>
<td>.058</td>
<td>-.075</td>
<td>.056</td>
</tr>
<tr>
<td>Scale 3: Hysteria</td>
<td>.061</td>
<td>-.182</td>
<td>-.037</td>
</tr>
<tr>
<td>Scale 4: Psychopathic Deviate</td>
<td>-.600</td>
<td>-.115</td>
<td>.244</td>
</tr>
<tr>
<td>Scale 5: Masculinity-Femininity</td>
<td>.080</td>
<td>.303</td>
<td>-.036</td>
</tr>
<tr>
<td>Scale 6: Paranoia</td>
<td>-.223</td>
<td>.259</td>
<td>-.267</td>
</tr>
<tr>
<td>Scale 7: Psychasthenia</td>
<td>.120</td>
<td>-.341</td>
<td>-.144</td>
</tr>
<tr>
<td>Scale 8: Schizophrenia</td>
<td>.162</td>
<td>-.026</td>
<td>.195</td>
</tr>
<tr>
<td>Scale 9: Hypomania</td>
<td>.052</td>
<td>.189</td>
<td>.114</td>
</tr>
<tr>
<td>Scale 0: Social Introversion</td>
<td>.271</td>
<td>.048</td>
<td>.210</td>
</tr>
</tbody>
</table>

Table 4.  
*Structure Coefficients for D-WESE Factor Scores with MMPI Clinical Scales*

<table>
<thead>
<tr>
<th>D-WESE Factors</th>
<th>Function 1</th>
<th>Function 2</th>
<th>Function 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive Functioning</td>
<td>-.793</td>
<td>.331</td>
<td>-.068</td>
</tr>
<tr>
<td>Anxiety Symptoms</td>
<td>-.766</td>
<td>.045</td>
<td>.040</td>
</tr>
<tr>
<td>Asocial Behaviors</td>
<td>-.326</td>
<td>.339</td>
<td>-.125</td>
</tr>
<tr>
<td>Depressive Features</td>
<td>-.716</td>
<td>-.633</td>
<td>.242</td>
</tr>
<tr>
<td>Somatic Symptoms</td>
<td>-.158</td>
<td>.034</td>
<td>.965</td>
</tr>
<tr>
<td>Inattention</td>
<td>-.750</td>
<td>.264</td>
<td>.032</td>
</tr>
</tbody>
</table>
Table 4 - Continued

<table>
<thead>
<tr>
<th>D-WESE Factors</th>
<th>Function 1</th>
<th>Function 2</th>
<th>Function 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychotic Symptoms</td>
<td>-.593</td>
<td>.499</td>
<td>-.137</td>
</tr>
<tr>
<td>Posttraumatic Symptoms</td>
<td>-.405</td>
<td>.121</td>
<td>.267</td>
</tr>
<tr>
<td>Energy Level</td>
<td>-.516</td>
<td>.085</td>
<td>.108</td>
</tr>
<tr>
<td>Appetite/Sleep</td>
<td>-.272</td>
<td>-.081</td>
<td>.450</td>
</tr>
<tr>
<td>Rebelliousness</td>
<td>-.165</td>
<td>.316</td>
<td>-.066</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D-WESE Factors</th>
<th>Function 4</th>
<th>Function 5</th>
<th>Function 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive Functioning</td>
<td>.027</td>
<td>.336</td>
<td>-.210</td>
</tr>
<tr>
<td>Anxiety Symptoms</td>
<td>.057</td>
<td>-.339</td>
<td>-.400</td>
</tr>
<tr>
<td>Asocial Behaviors</td>
<td>-.796</td>
<td>-.012</td>
<td>.298</td>
</tr>
<tr>
<td>Depressive Features</td>
<td>.067</td>
<td>.047</td>
<td>.020</td>
</tr>
<tr>
<td>Somatic Symptoms</td>
<td>.075</td>
<td>.070</td>
<td>-.059</td>
</tr>
<tr>
<td>Inattention</td>
<td>.127</td>
<td>-.109</td>
<td>-.048</td>
</tr>
<tr>
<td>Psychotic Symptoms</td>
<td>-.049</td>
<td>.246</td>
<td>.326</td>
</tr>
<tr>
<td>Posttraumatic Symptoms</td>
<td>.111</td>
<td>-.384</td>
<td>.543</td>
</tr>
<tr>
<td>Energy Level</td>
<td>.020</td>
<td>-.265</td>
<td>-.154</td>
</tr>
<tr>
<td>Appetite/Sleep</td>
<td>-.042</td>
<td>-.027</td>
<td>-.116</td>
</tr>
<tr>
<td>Rebelliousness</td>
<td>-.165</td>
<td>-.162</td>
<td>.250</td>
</tr>
</tbody>
</table>
The First Canonical Function

In the first canonical function, nine out of ten clinical scales of the MMPI had structure coefficients above .30 (Tabachnick & Fidell, 2001). Masculinity-Femininity was the only clinical scale that did not significantly contribute to the shared variance between the MMPI scales and the D-WESE factor scores in this function. Six of the ten scales, Depression, Psychopathic Deviate, Paranoia, Anxiety, Schizophrenia, and Hypomania, had structure coefficients above .50, suggesting a strong relationship (Cohen, 1992). Hypochondriasis, Hysteria, and Social Introversion had structure coefficients between .30 and .50, suggesting a moderate relationship (Cohen, 1992). These results indicate the majority of the content of the MMPI clinical scales contributed to shared variance between the MMPI and the D-WESE.

Eight of the eleven factor scores of the D-WESE substantially contributed to the canonical variable in the first function. Somatic Symptoms, Appetite/Sleep, and Rebelliousness did not significantly contribute to the canonical function. Intellectual/Executive Functioning, Anxiety Symptoms, Depressive Features, Inattention, Unstable Mood/Psychotic Symptoms, and Energy Level had structure coefficients above .50, suggesting a strong relationship (Cohen, 1992). Asocial Behaviors and Post-traumatic Symptoms had structure coefficients between .30 and .50, suggesting a moderate relationship (Cohen, 1992). Similar to the MMPI, the majority of the D-WESE factor scores contributed to the shared variance between the two measures.

To examine the extent to which the MMPI and the D-WESE account for common variability, redundancy coefficients for each were calculated. Redundancy coefficients indicate the proportion of variation in the set of one measure that is accounted for by the
other. Therefore, if the measures were found to have high redundancy coefficients, the results from one measure could be used to predict the results of the other. The MMPI redundancy coefficient for the first canonical function was .22. This suggests that approximately 22% of the variation in the D-WESE factors can be explained by the MMPI canonical variable. The D-WESE redundancy coefficient for the first canonical function was .29, suggesting that approximately 29% of the variation in the MMPI scales can be explained by the D-WESE canonical variable. These relatively low percentages indicate each measure is not highly predictive of the other.

The Second Canonical Function

In a second canonical function, only three of the ten scales of the MMPI had structure coefficients above .30. Depression, Hypomania, and Social Introversion had structural coefficients between .30 and .50. The analysis of the canonical variables for the D-WESE yielded five factors that significantly contributed to the second canonical function. Depressive Symptoms had a structural coefficient greater than .50. Intellectual/Executive Functioning, Asocial Behaviors, Unstable Mood/Psychotic Symptoms, and Rebelliousness had structural coefficients between .30 and .50.

As with the first canonical function, a redundancy analysis was conducted for the second canonical function. The MMPI redundancy coefficient was .06, while the D-WESE redundancy coefficient was .07 suggesting the canonical variables for each scale measured 6 and 7% of variance in the other scale respectively.

The Third Canonical Function

Three of the ten MMPI scales, Hypochondriasis, Depression, and Hysteria had structure coefficients above .30 in the third canonical function. These three scales make
up the neurotic triad on the MMPI (Greene, 2000). The D-WESE factors Somatic Symptoms and Appetite/Sleep significantly contributed to the third canonical function.

A redundancy analysis of the third canonical function yielded a coefficient of .07 for the MMPI and .09 for the D-WESE. As with the first two canonical functions, these values suggest a small degree of predictive value.

**The Fourth Canonical Function**

In the fourth canonical function, each measure yielded only one scale with a significant structure coefficient. The Psychopathic Deviate scale on the MMPI and the Asocial Behaviors factor on the D-WESE had structure coefficients above .30. Analysis of redundancy in the fourth canonical function yielded a coefficient of .02 for both the MMPI and the D-WESE.

**The Fifth Canonical Function**

The MMPI scales Masculinity-Femininity and Psychasthenia had structure coefficients above .30 in the fifth canonical function. The D-WESE factors Intellectual/Executive Functioning, Anxiety Symptoms, and Post-traumatic Symptoms had structure coefficients above .30 in this function. The redundancy analysis yielded a coefficient of .01 for the MMPI and less than .01 for the D-WESE.

**The Sixth Canonical Function**

In the sixth canonical function, none of the MMPI scales had structure coefficients above .30, indicating none of the MMPI scales contributed significantly to the variance accounted for in this function. The D-WESE factors Anxiety Symptoms, Unstable Mood/Psychotic Symptoms, and Post-traumatic Symptoms had structure
coefficients above .30 in the sixth canonical function. The redundancy analysis yielded a coefficient of .01 for the MMPI and less than .01 for the D-WESE.
CHAPTER 5
DISCUSSION

The present study examined the canonical relationship between well-established measures of emotional function and factors of psychiatric disorders based on DSM-IV-TR criteria. In contrast to past research on the subject that looked at the extent of the relationship between the MMPI and the D-WESE, the present study focused on how the constructs are related as well as how much the scales overlapped. Indeed, the purpose of the study was (a) to quantify the amount of shared variance between the MMPI scales and the D-WESE factors and (b) to calculate the amount of variance that could not be explained by the other measure. As hypothesized, the D-WESE factors were significantly correlated with the MMPI scales. This outcome suggests a substantial degree of relationship between psychiatric symptoms measured by a widely used objective psychiatric measure and a structured interview new to the field.

The Nature of the Relationship between the MMPI and the D-WESE

The current findings supported the recent trend in clinical diagnosis regarding the utility of structured diagnostic interviews in clinical practice. That is to say, though diagnostic interviews have been utilized since the inception of psychiatry, the use of
mutually agreed upon diagnostic criteria and structured interviews is a newer concept. One of the major limitations of unstructured interviews was the subjectivity with which diagnostic information was selected and gathered (Kamphaus & Frick, 2002). Structure diagnostic interviews such as the D-WESE minimize subjectivity by utilizing standard questions and responses that could be offered (Dean & Woodcock, 2003). The increased objectivity with which the information is gathered mirrors the methodology of personality inventories such as the MMPI. Therefore, a structured diagnostic interview and a personality inventory that purport to measure similar constructs would be expected to correlate significantly with one another. As previously mentioned, the D-WESE is a relatively new approach the field of neuropsychology. Best practice in assessment requires proof of psychometric properties of the assessment tool. Preliminary research supported an eleven factor structure as well as significant correlations between the MMPI and the D-WESE (Galloway-Sharp, 2004; Stage, in press). This study was rare in that it used a canonical analysis to look at the shared and unique variance between the MMPI and the D-WESE.

Canonical correlation is a multivariate procedure that focuses on a linear combination of variables that produces the largest correlation with a second set of variables. This linear combination is extracted and the process is repeated for the residual data, with the constraint that the second linear combination of variables must not correlate with the first one (Tabachnick & Fidell, 2001). The process is repeated until a successive linear combination is no longer significant (Tabachnick & Fidell, 2001). The results of the canonical correlation suggested a complex, significant relationship between the MMPI clinical scales and the D-WESE factor scores. The present results showed that
six of the ten canonical functions reached statistical significance. It is important to note that though the fourth, fifth, and sixth canonical functions were statistically significant, the percent of variance they explain is relatively low and provides less meaningful information than the first three functions. Canonical correlations are interpreted much like the Pearson r, in that a square of the measure provides the amount of variability accounted for.

In the first canonical function, 72.8% of the variance was shared between the MMPI clinical scales and the D-WESE factor scores. This function represents the strongest relationship between the two measures. Nine of the ten clinical scales of the MMPI contributed to this relationship. As predicted, Masculinity-Femininity did not significantly contribute to this canonical function (Greene, 2000). Hypochondriasis, Depression, Hysteria, Psychopathic Deviate, Paranoia, Psychasthenia, Schizophrenia, Hypomania, and Social Introversion scales each contained information that significantly related to content in the D-WESE factors.

The D-WESE factors that contributed to the first function included Intellectual/Executive Functioning, Anxiety Symptoms, Depressive Features, Inattention, Unstable Mood/Psychotic Symptoms, Energy Level, Asocial Behaviors and Post-traumatic Symptoms. These results were slightly different than expected. As discussed, the MMPI does not purport to assess Intellectual/Executive Functioning or Inattention, however, these two variables correlated significantly with MMPI clinical scales. Stage (in press) provided a possible reason for this shared variance. He found that Intellectual/Executive Functioning and Inattention factors contained items with multiple theoretical bases, suggesting they reflected more than one psychological construct.
Intellectual/Executive Functioning contained three items from different theoretical constructs, including psychotic symptoms, mood symptoms, and personality features, while Inattention contained three items associated with the theoretical construct mood symptoms (Stage, in press). The fact that so many of the MMPI clinical scales and D-WESE factor scores contributed to the shared variability of this first canonical correlation suggests the function represents an overall assessment of emotional functioning. If the results had not suggested this degree of overlap between the MMPI scales and the D-WESE factors, the validity of the D-WESE would be in question, as the lack of relationship would suggest the D-WESE does not measure psychiatric constructs important to psychiatric diagnosis and treatment.

In the second canonical function, 65% of the variance between the two measures was shared. The MMPI scales that contributed to this function, Depression, Hypomania, and Social Introversion, suggest a relationship based on mood symptoms. In fact, Depressive Symptoms had the strongest relationship to this function among D-WESE factors, and Unstable Mood/Psychotic Symptoms and Intellectual/Executive Functioning also significantly related. This supports the hypothesis that a strong relationship exists between MMPI scales 2 and 7 and D-WESE Depressive Symptoms and Unstable Mood/Psychotic Symptoms. Interestingly, the D-WESE factors of Asocial Behaviors and Rebelliousness also significantly contributed to this relationship. This relationship may suggest a link between asocial behaviors and lack of behavior control associated with manic diagnoses.

In the third canonical function, 57% of the variance was shared between the two measures. The MMPI Scales of Hypochondriasis, Depression, and Hysteria contributed
to this relationship. These three scales make up the neurotic triad, suggesting they measure internalizing emotional symptoms (Greene, 2000). The D-WESE factors Somatic Complaints and Appetite/Sleep were significant in this function, which supports the hypothesis that MMPI scale 2 would relate to D-WESE factor Appetite/Sleep, a measure of depressive symptoms. It also supports the hypothesis that MMPI scales 1 and 3 would related to the D-WESE factor of Somatic Complaints.

In the fourth canonical function, 26.7% of the variance between the two measures was shared. Both the MMPI clinical scale and the D-WESE factor score that significantly contribute to this function measure asocial behavior. These results reflect the hypothesis that both MMPI Scale 4 and the Asocial factor of the D-WESE measure a similar psychological constructs. The rebelliousness factor of the D-WESE was also hypothesized to relate to these scales, however, based on its relationship with mood scales, it seems to measure a different component of asocial behavior than MMPI scale 4 and the Asocial factor of the D-WESE.

In the fifth canonical function, 22.7% of the variance was shared between the two measures. Interestingly, in addition to the Psychasthenia scale, the Masculinity-Femininity scale contributed to this function. The D-WESE factors Intellectual/Executive Functioning, Anxiety Symptoms, and Post-traumatic Symptoms also significantly contributed, suggesting a relationship based on the construct of anxiety. It is unclear as to why the Masculinity-Femininity scale would contribute to a function related to anxiety. One explanation for this relationship may be the high degree of overlap of items across scales of the MMPI. For example, Scale 5 may share several items with Scale 7, causing a significant relationship between these variables.
In the sixth canonical function, 14.2% of the variance was shared. The MMPI scales did not significantly contribute to this relationship. The D-WESE factors of Anxiety Symptoms, Unstable Mood/Psychotic Symptoms, and Post-traumatic Symptoms were significant for this function. This function appears to account for mood and anxiety symptoms that did not contribute to the previous function. This function may reflect information gathered across scales of the MMPI that are more specifically categorized into the three significant D-WESE factors.

The strength of the canonical relationship between the MMPI and the D-WESE provided psychometric support for the use of the D-WESE, suggesting it measures relevant psychological constructs similar to those measured by the MMPI. In clinical practice, the D-WESE may offer several advantages over the MMPI. The items of the D-WESE evaluate both the presence and severity of symptoms in a structured manner that is organized along the DSM-IV diagnostic criteria. The results of the interview provide clear symptoms useful in diagnosis, treatment, and psychological theory (Dean & Woodcock, 2003). In contrast, the clinical scales of the MMPI provide behavioral correlates related to symptom groups (Greene, 2000). Therefore, information regarding certain symptoms, their duration and severity, and their interaction with treatment is lost. Thus, the MMPI may offer useful information in the overall assessment of emotional functioning, but it lacks information useful in understanding patient functioning. Unlike the MMPI, which requires hours to complete, the D-WESE requires about 20 minutes to complete, allowing for a broad range of information to be covered in a short period of time (Dean & Woodcock, 2003). The results of this study demonstrate validity for the D-WESE, providing stronger support for its use in clinical practice.
Unique Clinical Information Provided by the MMPI and the D-WESE

Recent research regarding the comorbidity of psychiatric features in neurological disorders suggests the need for current, accurate measures of emotional symptoms (Collins, Pastorek, Tharp, & Kent, 2012). Early in the field of neuropsychology, the MMPI was utilized to assess the degree to which patient disorders were more organic (biological) than functional (situational) (Reitan, 1955). The results of the data provided by the clinical scales and codetypes gave an overall picture of the patient’s profile and related behaviors. Current diagnosis has progressed from the concept of organic versus functional to a universally accepted set of specific symptoms and symptom groups provided by the *DSM-IV-TR*. The items of the D-WESE were written to reflect several diagnostic categories of the *DSM-IV-TR*, suggesting patient results may provide accurate information regarding patient emotional symptoms as specified by current diagnostic processes. The changes by which emotional symptoms are conceptualized and measured suggest a possible significant difference in the types of information collected by the MMPI and the D-WESE. In fact, despite the significant relationship between the MMPI clinical scales and the D-WESE factors, the redundancy analysis demonstrated that the measures accounted for a relatively small percentage of the variance of the other measure. A major question in this study was what percent of the D-WESE was accounted for by the MMPI scales and the percent of the D-WESE that accounted for the MMPI scales.

Redundancy is the amount of variance the canonical variates from one set of variables extract from the other and vice versa (Tabachnick & Fidell, 2001). High levels of redundancy suggest that the information gathered from one set of variables is very
similar to the information gathered by the other. Redundancy values for each canonical function can be summed to determine the overall extent to which one set of variables accounts for the variance in another set. A redundancy analysis of 100% would indicate that two sets of variables measure the same things and could be used interchangeably in practice. As mentioned, the redundancy analysis for this study demonstrated that the sets of variables for the MMPI and the D-WESE accounted for a relatively small percentage of the variance of each other. Across the six significant canonical functions, the MMPI clinical scales accounted for 39% of the variance of the D-WESE factor scores, while the D-WESE accounted for 48.5% of the variance of the MMPI clinical scales. Thus, 61% of the variance of the D-WESE is unique and 51.5% of the variance of the MMPI is unique. Therefore, although it was determined through the canonical functions that the MMPI clinical scales and the DWES factors share a significant amount of variance in terms of the information they measure, the redundancy analysis suggests the two instruments also measure a significant amount of different information. As hypothesized, Scale 5 of the MMPI did not significantly contribute to the primary relationship between the MMPI and the D-WESE. This may account for some of the unique variability in the MMPI. Results may also suggest the MMPI and the D-WESE measure the same psychological constructs in different ways that reflect the changes in the way psychiatric information is gathered. The value of the independent information measured by each instrument to current diagnostic processes remains unknown. Clinically, the results of this study indicate the MMPI and the D-WESE are not similar enough to be used interchangeably. Therefore, though the D-WESE may have multiple advantages in terms of ease of administration and information regarding duration and severity of psychiatric symptoms, it does not
cover all the information gathered by the MMPI. The results suggest practitioners may
benefit from administering both the MMPI and the D-WESE to gather the broadest range
of diagnostic information. However, future research is needed to determine whether one
of these measures provides information that is more helpful in current diagnostic
processes.

**Limitations**

Some limitations of the current study were noted regarding the sample. First, the
data sample was somewhat limited in size ($n = 207$) and therefore may be vulnerable to
limited representation and potential problems with homogeneity. Demographic
information for the sample set was limited to gender and age range. Additionally, the
data were gathered from a single geographical region. An archival data sample was used
due to limited availability of neuropsychological evaluations that included both the
MMPI and D-WESE, a relatively new measure. Future evaluations of the D-WESE may
benefit from developing a larger dataset with demographic heterogeneity in terms of
ethnicity and geographic region. In addition, the dataset utilized in this study was also
examined in two previous studies by Galloway-Sharp (2004) and Stage (in press), and the
results identified between these studies may therefore be influenced by potential
idiosyncrasies of this dataset.

This study also utilized the MMPI rather than the more recent revision of the
measure, the MMPI-2. However, the specific scales utilized in this study, the clinical
scales, share the vast majority of items between the MMPI and MMPI-2 (Levitt, 1990).
No datasets with adequate sample size including the MMPI-2 and the D-WESE were
available at the time of this study. Future studies should attempt to utilize the MMPI-2, a measure more commonly utilized in current research and practice.

**Future Directions**

One major limitation to the clinical use of the D-WESE is the lack of research regarding reliability of the measure. The current study focused on the validity of the D-WESE, specifically how well the D-WESE measures what it purports to measure. Future research to determine the reliability of D-WESE results over time and across raters is essential.

The results of the current study suggest that the MMPI and the D-WESE share a significant amount of variability in terms of what they purport to measure. However, the redundancy analysis indicates the two instruments account for a relatively small percentage of what the other measures, suggesting the instruments cannot be used to predict the outcome of the other. The question remains in regards to whether the MMPI or the D-WESE currently has stronger clinical utility. Future research using samples of clinical and nonclinical samples to determine which measure more accurately classifies these groups is warranted.

Research is also necessary regarding the ability of the MMPI and the D-WESE to determine co-morbidity, particularly with the co-occurrence of psychiatric and neurological symptoms. Difficulties in reaching accurate psychiatric diagnoses, specifically given problems related to overlapping scale items and a lack of specific diagnostic information on the MMPI, is compounded when more than one diagnosis is present. Unlike the MMPI, the D-WESE was developed as part of a neuropsychological battery and includes questions related to specific neuropsychological symptoms.
However, questions remain regarding the D-WESE’s ability to diagnose comorbid psychiatric and neurological conditions.

**Summary**

The MMPI and the D-WESE appear to measure similar information across several emotional constructs. However, both measures appear to gather a significant amount of information not accounted for by the other measure. The two measures gather information based on theoretical constructs relevant at the time of their development, with the MMPI focusing on organic versus functional explanations for neuropsychological symptoms and the D-WESE focusing on current symptoms and symptom groups as delineated in the *DSM-IV-TR*. More research is needed to determine whether one of these instruments is better than the other at predicting patient diagnosis. This information would allow clinicians to determine the utility of each measure in patient diagnosis and treatment.
REFERENCES


psychiatric comorbidities of TBI. In J. W. Tsao (Ed.) *Traumatic brain injury* (pp. 223-244). New York: Springer.


Freedland, K., Skala, J., Carney, R., Raczynski, J., Taylor, C., Mendes de Leon, C., Ironson, G., Youngblood, M., Krishnan, K., & Veith R. (2002). The Depression Interview and Structured Hamilton (DISH): Rationale, development,
characteristics, and clinical validity. *Psychosomatic Medicine, 64*(6), 897-905.


Griffiths, R. & Gillingham, P. (1978). The influence of videotape feedback on the self-
Assessments of psychiatric patients. *British Journal of Psychiatry, 133, 156-161.*


Manne, S., Nereo, N., DuHamel, K., Ostroff, J., Parsons, S., Martini, R., Williams, S.,


Miller, P. (2001). Inpatient diagnostic assessments: Interrater reliability and outcomes of
structured versus unstructured interviews. *Psychiatry Research, 105*(3), 265-271


