

THE CONCEPTUALIZATION OF PERSONALITY DISORDERS:
CATEGORIES VERSUS DIMENSIONS

Except where reference is made to the work of others, the work described in this dissertation is my own or was done in collaboration with my advisory committee. This dissertation does not include proprietary or classified information.

Danny R. Burgess

Certificate of Approval:

Chris J. Correia
Associate Professor
Psychology

Roger K. Blashfield, Chair
Professor
Psychology

Martha C. Escobar
Associate Professor
Psychology

Elizabeth B. Knight
Associate Professor
Psychology

Joe F. Pittman
Interim Dean
Graduate School

THE CONCEPTUALIZATION OF PERSONALITY DISORDERS:
CATEGORIES VERSUS DIMENSIONS

Danny R. Burgess

A Dissertation

Submitted to

the Graduate Faculty of

Auburn University

in Partial Fulfillment of the

Requirements for the

Degree of

Doctor of Philosophy

Auburn, Alabama
August 9, 2008

THE CONCEPTUALIZATION OF PERSONALITY DISORDERS:
CATEGORIES VERSUS DIMENSIONS

Danny R. Burgess

Permission is granted to Auburn University to make copies of this dissertation at its discretion, upon request of individuals or institutions and at their expense. The author reserves all publication rights.

Signature of Author

Date of Graduation

DISSERTATION ABSTRACT
THE CONCEPTUALIZATION OF PERSONALITY DISORDERS:
CATEGORIES VERSUS DIMENSIONS

Danny R. Burgess

Doctor of Philosophy, August 9, 2008
(M.S., Auburn University, 2004)
(B.S., University of Southern Mississippi, 2001)

104 Typed Pages

Directed by Roger K. Blashfield

The classification of mental disorders is organized as a categorical system, such that each disorder is a separate entity defined by an exclusive list of symptoms. Research has revealed many problems with the categorical conceptualization of mental disorders, and currently there is push for a dimensional system to replace the categorical framework, especially for personality disorders. In the current study, 115 psychologists from three different states described nine case vignettes using the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR)*, the Five Factor Model (FFM), and the Global Assessment of Functioning (GAF). The usefulness of the diagnostic systems, known as clinical utility, was measured through

self-report and behavioral procedures. Results show that the psychologists in this study were more familiar and confident using the *DSM-IV-TR*; however, they were able to employ the FFM and GAF systems better than the *DSM-IV-TR* when generating personality profiles of the case vignettes. In regards to behavioral evidence, participants utilized more FFM terminology when communicating information about the case vignettes. In light of the upcoming arrival of the *DSM-V* and the potential shift to a dimensional approach for personality disorders, more research is needed to examine how mental health professionals use different models of personality classification to improve communication, conceptualization, comprehensiveness, and treatment planning.

Style manual used: *Publication Manual of the American Psychological Association, Fifth Edition (2001)*

Computer software used: *Microsoft Word Release 10, Statistical Packages for the Social Sciences (SPSS) for Windows Release 12.0*

TABLE OF CONTENTS

| | |
|---|------|
| LIST OF TABLES | viii |
| INTRODUCTION | 01 |
| METHOD I | 15 |
| Participants..... | 15 |
| Materials and Procedures..... | 16 |
| METHOD II..... | 20 |
| Participants..... | 20 |
| Materials and Procedures..... | 21 |
| RESULTS | 23 |
| Participants..... | 23 |
| Combined Sample of Participants..... | 24 |
| Case Vignettes | 25 |
| Reliability Measure for the <i>DSM</i> and FFM Ratings | 30 |
| Clinical Utility Ratings | 30 |
| Comparing Across the <i>DSM</i> , FFM, and GAF Profiles | 32 |
| Word-List Worksheet..... | 35 |
| DISCUSSION..... | 38 |
| Comparisons Across the Classification Systems | 41 |
| Alex: A Case Example..... | 43 |
| Word-List Analysis..... | 47 |
| Limitations and Future Directions | 48 |
| Conclusion | 51 |
| REFERENCES | 53 |
| APPENDIX A: CASE VIGNETTES..... | 62 |
| APPENDIX B: EXAMPLES OF ALEX'S WORD LISTS..... | 80 |
| APPENDIX C: MATCHING TASK..... | 82 |

LIST OF TABLES

| | |
|--|----|
| 1. Description of Case Vignettes | 85 |
| 2. Percentage of <i>DSM</i> Diagnoses for Each Case Vignette..... | 86 |
| 3. Mean (<i>SD</i>) <i>DSM</i> Prototypical Ratings for Each Case Vignette..... | 87 |
| 4. Mean (<i>SD</i>) FFM Ratings for Each Case Vignette | 88 |
| 5. GAF Profiles for Each Case Vignette..... | 91 |
| 6. Reliability Coefficients for the <i>DSM</i> Prototypicality Ratings and FFM Facet Ratings..... | 92 |
| 7. Mean (<i>SD</i>) Utility Ratings of the <i>DSM</i> , FFM, and GAF for Each Case Vignette..... | 93 |
| 8. Mean (<i>SD</i>) Absolute Differences for Each Case Vignette Across the <i>DSM</i> (Based on the Mean of the Two Most Frequent Diagnoses), FFM (Based on the Mean of the Five Domains), and GAF (Based on Range) on a 7-Point Rating Scale | 94 |
| 9. Ranking of Variability Values for Each Case Vignette Across the Classification Systems | 95 |
| 10. Frequency (and Percentage) of Descriptors Across All Case Vignettes for Each Classification System..... | 96 |

INTRODUCTION

The current United States classification of mental disorders is found in the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR; American Psychiatric Association [APA], 2000)*. Each disorder is listed with its own criteria, and these disorders are arranged into broad groups based on similarity of symptoms, etiology, and behaviors. One such group is called the Personality Disorders, which are considered to be persistent, inflexible, maladaptive patterns of behavior resulting in significant distress or impairment in several areas of functioning (APA, 2000).

The *DSM-IV-TR* (APA, 2000) conceptualizes personality disorders as categories, suggesting that each disorder is its own unique entity with no overlapping criteria. Therefore, an individual is either given a personality disorder diagnosis or not, with no indication of to what degree an individual might represent that particular disorder. For example, schizoid personality disorder is defined by seven symptoms of which any four are sufficient for a diagnosis. Such a diagnostic label does not indicate if the individual is experiencing the minimum of four symptoms or all seven symptoms; nor does it suggest the degree of severity of the symptoms being expressed. Concerns have been raised regarding the validity of the *DSM* categorical approach (Ball, 2001; Clark, Livesley, & Morey, 1997; Widiger, 1992, 1993; Widiger & Francis, 2002; Widiger & Samuel, 2005), citing such problems as diagnostic overlap, heterogeneity within categories, arbitrary

boundaries for diagnosis, inadequate coverage, limited clinical utility, and loss of clinical information (i.e., degree of severity or impairment). The most common and well-established alternative to categories in classifying personality disorders is a dimensional approach. The Five Factor Model (FFM; Costa & McCrae, 1985) serves as one of the more familiar dimensional organizations of personality. This dimensional perspective places personality traits on continuums measuring the degree of intensity and severity. The FFM consists of five broad domains of personality functioning: Neuroticism, Extraversion, Openness to experience, Agreeableness, and Conscientiousness (Costa & McCrae, 1992). Each of these five domains is differentiated further into six specific facets. For example, the Neuroticism domain consists of anxiousness, angry hostility, depressiveness, self-consciousness, impulsivity, and vulnerability. The FFM provides a comprehensive description of both adaptive and maladaptive personality traits exhibited by an individual. Research has shown success with using the FFM as a diagnostic tool (Lynam & Widiger, 2001; Sprock, 2002; Widiger, 2000) and to guide treatment decisions and planning (Miller, 1991; Sanderson & Clarkin, 2002).

The historical roots of the FFM reach as far back as Galton (1884), one of the first researchers to examine the lexicon in describing meaningful personality differences through the use of single term descriptions. Thurstone (1934) took hold of this lexical hypothesis and through the use of factor analysis simplified a list of adjectives into five independent factors, concluding that "...the scientific description of personality may not be quite so hopelessly complex as it is sometimes thought to be" (p. 14). Exploration of the lexicon became more refined through the work of Allport and Odbert (1936), Cattell (1943, 1945), Eysenck (1947), Norman (1963, 1967), Guilford (1975), and Goldberg

(1976, 1982). Applying their strong factor-analytic backgrounds to dissecting the natural language, these researchers initiated the discussion of viewing personality differences through a dimensional lens, ultimately providing a strong foundation for future dimensional models to materialize. There has been much disagreement among these researchers on the ideal number of factors to sufficiently describe personality (typically ranging from three to over a dozen factors), but a number of investigators have shown that the five-factor structure emerges most consistently (Digman, 1990; Goldberg, 1990, 1992; McCrae & Costa, 1987; Wiggins & Pincus, 1992).

A second dimensional system used to describe psychopathology is the Global Assessment Functioning (GAF) scale. This one-dimensional approach is found within the *DSM-IV-TR* (APA, 2000) (on Axis V) and is used to measure the overall functioning of an individual. The GAF scale (scores range from 0 to 100) examines the level of functioning of the patient by evaluating the effects of their symptoms on psychological, social, and occupational domains.

The current GAF scale has gone through an extensive evolution, tracing back to Luborsky's (1962) Health-Sickness Rating Scale (HSRS). This 0-100 point scale offered a single rating to describe the healthiness (or sickness) of a patient on a variety of domains, including autonomy, use of abilities, interpersonal relationships, and levels of interests. The HSRS further proposed recommendations of care and treatment placement, such that particular ratings dictated that patients should be hospitalized or receive a specific intervention. A modified version of the HSRS appeared about 15 years later, known as the Global Assessment Scale (GAS; Endicott, Spitzer, Fleiss, & Cohen, 1976). The GAS differed in many ways from the HSRS, most notably with the removal of

mental health service suggestions (given the significant changes in healthcare and treatments offered) and dividing the 100-point scale into 10-point increments.

Despite the establishment of the GAS, the advent of the multi-axial psychiatric diagnostic system in the *DSM-III* (APA, 1980) gave rise to a different scale as a measure of functioning. Axis V, of the five-axis system, debuted in *DSM-III* as a 0 to 7 scale to measure an individual's highest level of functioning during the past year. Ratings were based on the individual's quality of social relations, occupational functioning, and use of leisure time. The establishment of the multi-axial system, including Axis V, generated little research (Skodol, Link, Shrout, & Horwath, 1988); therefore, Axis V was replaced by the GAS in the *DSM-III-R* (APA, 1987), given the clinical and research success of the GAS (Goldman, Skodol, & Lave, 1992; Skodol et al., 1988). The GAS was renamed the Global Assessment of Functioning and was based on social, occupational, and psychological functioning. The properties of the scale changed to a 1 to 90 scale, eliminating the 91-100 range of the GAS as it was thought individuals in this range likely would not seek mental health services. *DSM-IV* (APA, 1994) established the currently existing GAF scale, with the only major change from *DSM-III-R* being the inclusion of the 91-100 range of functioning. Little research exists that focuses exclusively on the clinical usefulness and properties of the GAF, as Axis V tends to be overshadowed by the clinical and research dominance of Axes I and II (Moos, McCoy, & Moos, 2000).

Criticisms about the GAF have emerged suggesting that a single GAF value is not sufficient in describing an individual's psychological, social, and occupational functioning (Goldman, Skodol, & Lave, 1992). The *DSM-IV-TR* (APA, 2000) included in the appendix two proposed axes in need of further research that would isolate an

individual's relationship functioning (Global Assessment of Relational Functioning [GARF] Scale) and an individual's social and occupational functioning (Social and Occupational Functioning Assessment Scale [SOFAS]). The GARF and SOFAS offer separate, and more specific, measurements for an individual's relational functioning and social and occupational functioning, respectively. On the other hand, a single GAF score attempts to represent the combination of all of these areas, as well as psychological impairment. Some research has begun to compare reliability and validity properties among the GAF, GARF, and SOFAS (Hay, Katsikitis, Begg, Da Costa, & Blumenfeld, 2003; Hilsenroth et al., 2000); however, Axis V continues to be neglected due to the emphasis placed on Axes I and II.

The debate between categorical versus dimensional models remains an active issue concerning the classification of personality disorders (Clark, 2007; Trull & Durrett, 2005) and the future editions of the *DSM* (Widiger & Samuel, 2005). Such recent discussions and debates were featured in a separate section of the *Journal of Abnormal Psychology* (Watson, 2005) and a special issue of the *Journal of Personality Disorders* (Livesley, 2005), which highlighted the implications of replacing categories with dimensions on such agendas as research, diagnostics, and clinical utility. Kraemer, Noda, and O'Hara (2004) argued that one model is not superior to the other but are equivalent, and both should be available to choose from, depending on the clinical or research circumstances. Other researchers have argued that the current *DSM* personality disorders can be better understood and described using the dimensional characteristics of the FFM (Lynam & Widiger, 2001; Samuel & Widiger, 2004; Saulsman & Page, 2004; Trull, 1992).

Despite these claims, the pure categorical approach continues to remain the standard form of classifying psychopathology in clinical and research settings. The categorical model has persisted in spite of little empirical evidence to support Axis II cluster structures and individual personality disorder criteria (Sheets & Craighead, 2007). However, the research on the dimensional approach is gaining momentum, suggesting that a dimensional conceptualization may be viewed by clinicians as a more useful and descriptive classification for mental disorders (Samuel & Widiger, 2006). Bernstein et al. (2007) surveyed members of the Association for Research on Personality Disorders (ARPD) and the International Society for the Study of Personality Disorders (ISSPD) to solicit expert opinions on the future of the classification of personality disorders. Eighty percent of the participants reported that their personal conceptualization of personality disorders are best represented as dimensions of psychopathology and over 85% of the experts reported the existing research data in the literature suggest the same. Furthermore, 56% of the participants are in favor of a dimensional system for Axis II disorders for the *DSM-V*, with almost 70% hoping for a mixed system of categories and dimensions. Only a fourth desires no change in the *DSM-V* from the current categorical system for personality disorders. Clearly, the majority of experts of the personality disorders are dissatisfied with the current conceptualization of the Axis II classification system. However, one noteworthy limitation of this survey study is response bias, as those experts most dissatisfied with the current classification of personality disorders may have been more motivated to participate in the survey.

There is some uncertainty of whether clinicians and researchers would embrace such drastic changes to the current classification system, or if these changes would create

even more problems than what currently exists with the *DSM* approach. Over 80% of the experts in the Bernstein and colleagues' (2007) study reported the current categorical classification is helpful with communication, and almost two-thirds described Axis II as easy to use. However, ease of use and communication should not be translated as a classification system that is valid, precise, and comprehensive (Clark, 1993). First et al. (2004) contended that more emphasis has been placed on diagnostic validity and reliability when considering the change from categories to dimensions to represent psychopathology, with little empirical investigation committed to examining clinical utility. For such a drastic change to occur in classifying mental disorders, more efforts are needed in determining whether or not clinicians and researchers would find a new system useful and as an improvement from the current classification system. First et al. proposed a formal definition of clinical utility that suggested a classification should aid in conceptualizing diagnoses, communicating such information to patients and professionals, differentiating diagnoses from other categories or criteria sets, developing a treatment intervention, and predicting future clinical needs. Using such a definition, the authors called for the employment of empirical methods to examine clinical utility for any proposed changes to the current classification system.

Clinical utility has been an intended priority of the modern *DSMs*; however, attention to clinical utility has just begun to drastically increase given the anticipated arrival of the *DSM-V* and the possible transition to a dimensional system for personality disorders. Various *DSM* work groups and international conferences have sponsored research agendas across a wide range of topics in preparation for the *DSM-V* (Widiger, Simonsen, Krueger, Livesley, & Verheul, 2005). One such topic has focused on obtaining

empirical data on the comparison of clinical utility between categorical and dimensional models. Verheul (2005) concluded that there is little evidence to support strong clinical utility when using a categorical or dimensional model, specifically in regards to coverage, reliability, subtlety, and treatment planning. Despite such prominence placed on clinical utility, only a few studies (Blais, 1997, Samuel & Widiger, 2004, 2006; Sprock, 2002, 2003) have empirically examined how psychologists apply different classification systems in clinical practice. Two of these clinical utility studies served as models for this current study.

Sprock (2003) compared the *DSM-IV* (APA, 1994) categorical system with a variety of dimensional and hybrid systems in diagnosing personality disorder case vignettes. She further examined the participants' perceived clinical utility of the classification systems. Practicing psychologists from the American Psychological Association were recruited to read three of six case vignettes depicting various personality disorders. One group of the sample utilized the *DSM-IV* categorical approach and two hybrid models (a mixture of dimensional ratings and personality disorder categories/clusters). The second group was required to use the FFM (McCrae & Costa, 1990), the seven-factor model (Cloninger & Svrakic, 1994), Siever and Davis' (1991) four-factor model, and the interpersonal circumplex as described by Wiggins and Pincus (1994). All participants were instructed to apply their respective classification model to the case vignettes and then make clinical utility ratings on their confidence with the model and its usefulness in communication, treatment planning, and case conceptualization.

The case vignettes represented each of the three personality disorder clusters as classified by the *DSM-IV* (APA, 1994), and half of the cases were considered prototypic (at least 80% diagnostic agreement) and the remaining cases were nonprototypic (less than 60% agreement). Participants received one case vignette (either prototypic or nonprototypic) from each personality cluster and provided diagnostic profiles using their respective classification models. The *DSM-IV* categorical approach was rated higher on confidence in using that system and on the clinical utility domains as compared to the hybrid models and the pure dimensional models. Furthermore, the hybrid models received higher confidence ratings and clinical utility ratings than the dimensional models. Of the dimensional models, Siever and Davis' (1991) four-factor model received the highest ratings for the clinical utility domains. These trends held despite the degree of prototypicality of the case.

Sprock (2003) concluded that even though the dimensional models were reliable in identifying and describing the pathology of the cases, clinicians did not find these models to be as useful, nor did they feel as confident employing these models as compared to the categorical approach or the hybrid models. Sprock suggested that if a transition from categories to dimensions were to occur for personality disorders, a hybrid model appears to have strong reliability and clinical utility properties to support such a conversion (see also Skodol et al., 2005). However, Sprock recognized the limitation of the case vignettes used in the study as not being representative of real-life individuals with personality problems. She suggested that dimensional models may outperform a categorical approach when a clinician is faced with an actual patient presenting with an extensive and complex history of symptoms, behaviors, relationships and functioning.

Samuel and Widiger (2006) continued the work of Sprock (2003) and instructed clinicians to employ the *DSM* and FFM classification systems to case vignettes and rate the clinical utility of each model. Samuel and Widiger recruited psychologists from Division 42 (Psychologists in Independent Practice) of the American Psychological Association and asked them to provide *DSM* and FFM ratings for three case vignettes that depicted real-life individuals with significant personality impairment. One vignette was a description of Ted Bundy, a notorious serial killer who used his charm and intelligence to entice young, college women with the goal to rape and murder them. The second vignette was Earnst, a patient of the Harvard Psychological Clinic who was made famous by appearing in Henry A. Murray's 1938 text *Explorations in Personality*. Earnst's troubled childhood led him to a life of mediocrity and loneliness, and he eventually sought refuge in a fantasy world of storybook characters. Earnst was unsuccessful in securing a good job or maintaining stable relationships due to his personal limitations. The third case vignette was Madeline, who appeared as a case study in Jerry Wiggins' 2003 text *Paradigms of Personality Assessment*. Madeline was a Native American who suffered through an abusive childhood. As an adult, she engaged in self-destructive behaviors through alcohol abuse and criminal activities. However, she improved her life and eventually established a successful law firm for a brief period of time until more personal problems emerged, resulting in the loss of her legal practice.

Psychologists were randomly given one case vignette and asked to determine FFM ratings on all 30 facets and *DSM* ratings for all ten personality disorders using a 5-point Likert scale. The sample was further instructed to make a final *DSM* diagnosis by selecting one or more of the personality disorders or no personality disorder. Once

profiles were generated for the *DSM* and FFM, participants completed a clinical utility questionnaire that assessed the ease of applying each system, usefulness in communication (with the individual and with another professional), comprehensiveness in describing personality problems, helpfulness in formulating a treatment plan, and comprehensiveness in globally describing personality.

The authors expressed that the primary goal of their investigation was to compare the utility ratings between the *DSM* and FFM. Analyses on the clinical utility ratings suggested that the psychologists agreed that the *DSM* and FFM were both moderately useful in the ease of applying each system to the case vignettes. However, the FFM was rated significantly higher than the *DSM* in usefulness in communicating the personality impairments to the individual, comprehensiveness in describing personality problems, helpfulness in formulating treatment plans, and comprehensiveness in globally describing personality for all three cases.

The authors concluded that the FFM had greater clinical utility than the *DSM*, which is inconsistent with the findings of Sprock (2003). However, Samuel and Widiger (2006) noted that Sprock's use of case vignettes may have influenced the results due to the brevity of the vignettes and the wording of the vignettes, as they were primarily based on *DSM* diagnostic criteria, thus possibly increasing the ease of using the *DSM* system. Samuel and Widiger contended that the FFM may serve as a more useful classification system when evaluating the complexity of real-life personality presentations. In regards to clinical utility, the problem with the Sprock and Samuel and Widiger studies was their sole use of a rating scale to measure how useful the clinicians found the different classification models to be. A more compelling argument would be to examine what

clinicians actually do with the classification systems rather than simply providing subjective ratings.

The current project expanded on Samuel and Widiger's (2006) research by including a third classification system (the GAF) to serve as the control, more case vignettes covering a wider range of psychopathology, and a behavioral approach to measure clinical utility. The first goal of the research study was to replicate the methodology of Samuel and Widiger by having participants apply the *DSM* and FFM systems to the case vignettes of Ted, Earnst, and Madeline. The use of case vignettes was expanded by adding six more to be assessed. All nine cases depicted either real-life individuals or characters from the popular media, thus providing richer and more complex clinical information that is not necessarily biased towards a particular classification system. Furthermore, the addition of more case vignettes offered a wider range of personality psychopathology to be evaluated, inevitably exposing psychologists to varying degrees of difficulty and familiarity in regards to conceptualizing personality impairment.

In addition to including more case vignettes, an additional classification system was used to serve as an alternative to the *DSM* and FFM. Samuel and Widiger (2006) showed, in general, that clinicians preferred the comprehensiveness and usefulness of the FFM over the *DSM*. However, this does not imply that their sample preferred the FFM. What if the clinicians disliked the *DSM* so much that any classification system would have been rated higher? If this was the case, then providing only one other option (the FFM) guaranteed a higher rater in clinical utility, even though the FFM may not ultimately be the desired system. Therefore, a third classification model was needed to

serve as a control variable. The GAF was chosen for the current study due to its familiarity to psychologists and its establishment on the multi-axial diagnostic structure (as opposed to the GARF or SOFAS). The GAF was also an ideal alternative because of its lack of sophistication and minimal research support, relative to the *DSM* and FFM. It was expected that psychologists will at least find the more sophisticated and comprehensive systems of *DSM* or the FFM to be more clinically useful than a simple dimensional scale such as the GAF.

To broaden the work of Samuel and Widiger (2006), a different approach was used for this study to provide behavioral evidence for clinical utility, in addition to self-report evidence. Participants completed a task in which they described from memory each case vignette using one to three word descriptors. Because self-report ratings of clinical utility do not adequately capture the usefulness or ease of a classification system, behavioral evidence of a clinician actually using the system to describe, communicate, or treat an individual was needed. As Goldberg aptly stated, "...the 'importance' of an individual difference is given operational definition as its probability of occurrence in the natural language" (1972, p. 548). Therefore, the responses on the memory task provided such evidence in the form of the psychologists' use of terminology associated with a particular classification system to conceptualize and communicate details about the case vignettes.

With the addition of more case vignettes, an alternative classification system to the *DSM* and FFM, and a different approach to assess clinical utility, the current study extended upon the work of Samuel and Widiger (2006) and may potentially contribute to the debate on how personality impairment should be classified. This study will build on

the emerging research thread of evaluating which classification system is the most useful, precise, and comprehensive way of conceptualizing personality presentations, in order to enhance communication, treatment planning, and description of clinical profiles.

METHOD I

Participants

A roster of licensed psychologists was obtained from the Alabama Board of Examiners in Psychology. The roster contained the name, mailing information, and license status of all psychologists who are licensed in the state of Alabama. As of October 2005, the roster listed 812 licensed psychologists. Two psychologists were removed from the list because of their probation status. Eleven other psychologists were removed because of their affiliation with the clinical psychology graduate program at Auburn University. Therefore, 799 licensed psychologists remained in the final list of potential participants. An initial sample of 400 psychologists was randomly chosen for the current study. The psychologists were contacted by mail to participate. The psychologists received an Information Sheet approved by the Institutional Review Board (IRB), which explained the purpose and procedure of the study, and a self-addressed, stamped postcard. The psychologists were instructed to send back the postcard if they agreed to participate. Upon receiving their postcard, the principle investigator mailed a research packet to the participants requesting them to complete the packet and mail it back.

Due to a low response rate during the first sampling procedure, a second sampling was conducted to include the remaining 399 licensed psychologists. The procedure remained the same as described above, except an additional letter was sent with the IRB Information Sheet and the postcard. The content of the letter was more personal in nature

than the standard IRB letter, and it described the importance of their contribution to the debate on the conceptualization of personality disorders. It was believed that the personal letter would increase the response rate, as the psychologists may respond more positively to the informal letter.

Materials and Procedure

The research packets contained three randomly chosen case vignettes out of nine total cases. Furthermore, the packet included guides for the *DSM-IV-TR* (APA, 2000) personality disorders, the FFM, and the GAF. Participants received three diagnostic profile score sheets for each case vignette for the *DSM*, the FFM, and the GAF, totaling nine score sheets. Finally, the research packet contained a word-list worksheet and a demographic questionnaire. Participants were provided with a self-addressed, stamped envelope to mail back the completed materials.

Case vignettes. Nine case vignettes were generated that highlighted individuals exemplifying varying degrees of psychopathology, dysfunction, and distress. The case vignettes were based on historical and popular media characters and were typically 1.5 to 1.75 pages long, single spaced. Three of the case vignettes (Ted, Earnst, Madeline) were generated by Samuel and Widiger (2006). The remaining six case vignettes were generated by the principal investigator and the faculty supervisor of the study. Table 1 provides brief descriptions and the length (in terms of word and sentence count) for each of the nine case vignettes, and Appendix A contains each case vignette in its entirety.

DSM, FFM, and GAF guides. Participants also received brief guides that outlined the classification of personality disorders found in the *DSM-IV-TR* (APA, 2000), the personality domains and facets of the FFM, and the levels of functioning of the GAF (as

described in the *DSM-IV-TR*). The guide for the *DSM* contained the general diagnostic criteria for a personality disorder and the criteria for each of the 11 personality disorders, including personality disorder Not Otherwise Specified (NOS). The guide for the FFM contained a general discussion about the development and use of the FFM, as well as a description of each of the five domains and their respective facets. The FFM guide was generated by Samuel and Widiger (2006); however, the author of the current study created a shorter, revised version by reducing the paragraph descriptions of each facet to a list of adjectives describing the extreme poles of each facet. Both the *DSM* guide and the FFM guide were six pages each (single space). The one-page GAF guide consisted of a brief description on the purpose and use of the GAF scale and a condensed version of the GAF scale, based on the GAF found in the *DSM-IV-TR*.

DSM score sheets. The *DSM* score sheet asked that the participants select the personality disorder(s) the patient met criteria. All 11 personality disorders were available to choose from, and the participants were allowed to choose as many personality disorders as they deemed appropriate. Furthermore, participants were instructed to make a prototypicality rating on a 7-point Likert scale (1 = not at all prototypical; 4 = somewhat prototypical; 7 = prototypical) for each of the personality disorder categories. Participants made ratings for all categories, regardless if they diagnosed the patient with that personality disorder. Finally, the score sheet contained two questions assessing the participant's confidence in the patient's *DSM* profile on a 7-point Likert scale (1 = not at all confident; 4 = somewhat confident; 7 = very confident) and the usefulness of the *DSM* system in describing the patient (1 = not at all useful; 4 = somewhat useful; 7 = very useful).

FFM score sheets. The FFM score sheet contained ratings scales for all 30 facets. Participants were instructed to make a rating on a 7-point Likert scale (1 = problematic, very low on the trait; 4 = neither high nor low on the trait; 7 = problematic, very high on the trait) for each facet based on the degree to which the patient exemplified that particular trait. On the score sheet, examples were provided to describe the extreme poles of each trait. For example, the facet assertiveness (found under the domain of Extraversion) ranged from unassuming, quiet, and resigned (a possible rating of 1 or 2) to dominant and forceful (a possible rating of 6 or 7). The FFM score sheet also included two questions that assessed the participant's confidence in the patient's FFM profile and the usefulness of the FFM system in describing the patient.

GAF score sheets. Participants were asked to provide the most accurate GAF score for the patient, as well as the highest and lowest possible GAF scores. As with the *DSM* and FFM score sheets, the GAF score sheet assessed the participant's confidence in the patient's GAF profile and the usefulness of the GAF system in describing the patient. The GAF score sheet also contained an opportunity for the participant to attempt to identify the patient, as the patients in the case vignettes were historical figures and popular media characters.

Word-list worksheet. Participants were asked to complete the word-list worksheet once they finished all other packet materials. They were requested to generate a list of brief descriptions (one to three words) that best characterized each of the three case vignettes they reviewed. The participants were instructed not to refer back to the case vignettes or the score sheets but perform this task strictly from memory. The word-list

worksheet examined how the psychologists conceptualized the case vignettes by their use of *DSM*, *FFM*, or *GAF* terminology.

Demographic questionnaire. All participants were asked to complete a demographic questionnaire. Information was gathered on basic characteristics of the sample (i.e., age, sex, degree, years of clinical experience, specialty, certification), as well as their familiarity with the *DSM*, *FFM*, and *GAF* systems based on a 7-point Likert scale (1 = not at all familiar; 7 = very familiar). Also, the questionnaire allowed participants to list what assessment tools they use when conducting personality assessments and evaluations.

METHOD II

Participants

Idaho sample. A current roster of licensed psychologists was obtained from the Idaho State Board of Psychologist Examiners. The roster contained the name, mailing information, and license status of all psychologists who are licensed in Idaho. As of November 2006, the Idaho roster listed 304 active licensed psychologists. All Idaho psychologists were contacted by mail to participate.

Connecticut sample. A current roster of licensed psychologists was obtained from the Connecticut Board of Examiners of Psychologists. The roster contained the name, mailing information, and license status of all psychologists who are licensed in Connecticut. As of March 2007, the Connecticut roster listed 1658 active licensed psychologists, and 825 were randomly chosen to participate.

Both samples of psychologists received an Information Sheet approved by the Institutional Review Board (IRB), which explained the purpose and procedure of the study, a personal letter describing the importance of their contribution to the debate on the conceptualization of personality disorders, and a self-addressed, stamped postcard. The psychologists were instructed to send back the postcard if they agreed to participate. Upon receiving their postcard, the principle investigator mailed a research packet to the participants requesting them to complete the packet and mail it back.

Materials and Procedure

All the materials (case vignettes, classification guides, score sheets, word-list worksheet, and demographic questionnaire) for Method II were the same from Method I. However, some procedural changes were made from Method I to Method II.

First, participants were instructed to complete the word-list task at the beginning of the study rather than at the end (as was described in Method I). The rationale for this change was the possibility that reading the classification manuals and responding to the score sheets may prime participants to utilize terminology they would not naturally use to describe the case vignettes.

Second, participants of Method II were given a note card (contained in the research packet), and they were instructed to provide their name, address, and any memory cues that would help them to remember the case vignettes. Once the participants completed the research packet, they mailed back their responses, with the note card included, in the provided self-addressed, stamped envelope. A second word-list worksheet (along with the note card containing their memory cues) was mailed to the participants, and they were asked to perform the word-list task again. Upon their completion of the second word-list task, they mailed back their list in the provided self-addressed, stamped envelope. Responses on the first word-list worksheet, the note card, and the second word-list worksheet were compared to see if the terminology of the three classification systems was used differently throughout the three stages of the word-list task (i.e., immediate response to the case vignettes vs. memory aides vs. delayed recall of the case vignettes).

Finally, a small subset of the Idaho and Connecticut samples ($n = 24$) received an additional task in which participants were required to match predetermined diagnostic profiles (for each of the three classification systems) to a set of case vignettes. Those participants who received the matching task were only given two case vignettes (as opposed to three) to generate *DSM*, *FFM*, and *GAF* diagnostic profiles. Due to low response rates, the matching task was eliminated from further data collection attempts. Presumably, participants may have found the study to be too lengthy and demanding of their time, thus resulting in such a low return of completed packets. Preliminary analyses and implications of the matching task will be discussed in a later section.

RESULTS

Participants

Alabama sample. Of the 799 request letters that were sent to the licensed psychologists, nine letters were returned due to changes in addresses, leaving a total of 790 possible participants. One hundred and forty-one (17.85%) postcards were returned indicating consent to participate in the study. Therefore, 141 research packets were mailed out and 71 (50.35%) were returned with usable data. In terms of all Alabama licensed psychologists that were recruited to participate ($N = 790$), this study represented 8.99% ($n = 71$) of this group.

Idaho sample. Of the 304 request letters that were sent to the licensed psychologists, 11 letters were returned due to changes in addresses, leaving a total of 293 possible participants. Fifty (17.06%) postcards were returned indicating consent to participate in the study. Therefore, 50 research packets were mailed out and 13 (26.00%) were returned. In terms of all Idaho licensed psychologists that were recruited to participate ($N = 293$), this study represented 4.44% ($n = 13$) of this group.

Connecticut Sample. Of the 825 request letters that were sent to the licensed psychologists, seven letters were returned due to changes in addresses, leaving a total of 818 possible participants. One hundred and one (12.35%) postcards were returned indicating consent to participate in the study. Therefore, 101 research packets were mailed out and 31 (30.69%) were returned. In terms of all Connecticut licensed

psychologists that were recruited to participate ($N = 818$), this study represented 3.79% ($n = 31$) of this group.

Combined Sample of Participants

Due to low response rates from the samples of Idaho and Connecticut, data from these groups were combined with the Alabama sample. All further analyses described here are based on this combined dataset.

Of the 1928 total request letters that were sent to the licensed psychologists for all three states, 27 letters were returned due to changes in addresses, leaving a total of 1901 possible participants. Two hundred ninety-two (15.36%) postcards were returned indicating consent to participate in the study. Therefore, 292 research packets were mailed out and 115 (39.38%) were returned. In terms of all licensed psychologists for the three states that were recruited to participate ($N = 1901$), this study represented 6.05% ($n = 115$) of this group.

The combined sample consisted of 50.00% females. The mean age was 47.59 years ($SD = 10.42$, $Mdn = 47.00$), ranging from 28 to 72 years of age. The majority of the participants were PhD psychologists (84.21%), with the remaining sample comprised of PsyD (11.40%), Master's level (2.63%), and EdD (1.75%). On average, the psychologists had 14.49 years of clinical experience ($SD = 9.97$, $Mdn = 12.00$), ranging from 0 to 36 years.

In terms of personality evaluation and assessment, 77.19% of the participants indicated they use the Minnesota Multiphasic Personality Inventory—Second Edition (MMPI-2), 35.96% use the Rorschach, just over a third reported the Thematic Apperception Test (TAT) as a part of their assessment, a fourth use the Personality

Assessment Inventory (PAI), 24.56% utilize some version of the Million inventories (i.e., MCMI—III, MACI), and 7.02% indicated the NEO-Personality Inventory-Revised (NEO-PI-R) is used in personality assessment.

Participants were also asked to rate their familiarity with the *DSM*, FFM, and GAF on a 7-point Likert scale (1 = not at all familiar; 4 = somewhat familiar; 7 = very familiar). The average familiarity ratings for the *DSM* and FFM were 6.05 ($SD = 0.83$) and 2.98 ($SD = 1.71$), respectively. For the GAF, the participants rated their familiarity as 5.69 ($SD = 1.09$). A one-way, repeated measures ANOVA found significant mean differences in familiarity among the systems, $F(2, 226) = 252.32, p < .001$. Pairwise comparisons revealed that the familiarity for the *DSM* was significantly higher than the FFM, $t(113) = 19.01, p < .001$, and the GAF, $t(113) = 4.11, p < .001$. Furthermore, the psychologists reported the GAF as being more familiar than the FFM, $t(113) = 14.86, p < .001$. Samuel and Widiger (2006) reported similar familiarity ratings for the *DSM* and FFM. Participants in their study clearly were more familiar with the *DSM* ($M = 4.26$, out of a 5-point scale) compared to the FFM ($M = 1.93$).

Case Vignettes

DSM profiles. For each case vignette, percentages of diagnosed personality disorders (refer to Table 2) and mean prototypicality ratings were determined (refer to Table 3). All case vignettes (except for Marianne and Macon) obtained a majority consensus (greater than 50%) for a particular personality disorder. Few cases reached overwhelmingly high levels of agreement (over 75%) on a personality disorder diagnosis (Alex for borderline; Ted and Aileen for antisocial). The minimum number of diagnoses given was five (Madeline), and the maximum was ten diagnoses (Charles).

A vast majority of the psychologists diagnosed Ted with antisocial personality disorder (83.9%), with the second leading diagnosis being narcissistic personality disorder (48.4%). For Madeline, just over 60% of the psychologists decided she met criteria for histrionic personality disorder, and over 50% indicated a diagnosis of narcissistic personality disorder. Just over two-thirds of the psychologists labeled Earnst as avoidant personality disorder, and 53.6% of the sample diagnosed him with schizoid personality disorder. The diagnoses given to these individuals corresponded with the findings of Samuel and Widiger (2006); however, the aforementioned authors reported greater percentage values. Samuel and Widiger reported that 96% of their sample diagnosed Ted with antisocial personality disorder and almost as many psychologists provided Ted with a diagnosis of narcissistic personality disorder. In regards to Madeline, Samuel and Widiger reported that 91% provided a diagnosis of narcissistic personality disorder and 87% diagnosed her with histrionic personality disorder. Madeline was also given a diagnosis of borderline personality disorder by two-thirds of Samuel and Widiger's sample, which was not a consistent finding in the current study. For Earnst, Samuel and Widiger reported 94% provided a diagnosis of avoidant personality disorder, and 80% diagnosed him with schizoid personality disorder.

Charles was given every personality disorder diagnosis except for dependent personality disorder. The majority of the sample diagnosed him with schizotypal personality disorder (58.6%) and narcissistic personality disorder (56.7%). A large consensus was reached with Aileen, evidenced by over 80% of the psychologists describing her as having antisocial personality disorder. Also, 48.6% of the psychologists provided a borderline personality disorder diagnosis for Aileen. A majority of the

psychologists perceived Meursault as having schizoid personality disorder (53.3%), with the second leading diagnosis being personality disorder NOS (30.0%).

Psychologists seemed to struggle the most with Marianne and were unable to reach a consensus on a specific personality disorder. Consequently, half of the participants provided a personality disorder NOS; notably, no clinician indicated that she was without a personality disorder. In other words, every participant diagnosed Marianne with a personality disorder, whereas all the other case vignettes received at least one report that they did not meet criteria for a personality disorder.

The highest agreement among the psychologists occurred with Alex, who received a borderline personality disorder diagnosis by over 90% of the participants. Conversely, Macon seemed to bewilder the psychologists, because a majority consensus was not reached. Just over a third of the psychologists considered him to have obsessive-compulsive personality disorder and just under a third diagnosed him with schizoid personality disorder.

Generally speaking, it was evident that the majority of the case vignettes exemplified more Cluster B personality disorders (antisocial, borderline, histrionic, narcissistic), which tend to be the dramatic and impulsive personality presentations. What was under-represented were the Cluster C personality disorders (avoidant, dependent, obsessive-compulsive), where patients present with a more fearful and anxious demeanor. When strictly looking at the top two diagnoses for each case vignette, Cluster B personality disorders made up 56% of the diagnoses, Cluster A personality disorders (paranoid, schizoid, schizotypal) consisted of 22% of the diagnoses given, and Cluster C and NOS each represented 11% of the chosen diagnoses.

In terms of prototypicality ratings, the participants agreed that Ted and Aileen were prototypical cases of antisocial personality disorder and Alex was a prototype of borderline personality disorder (mean ratings greater than 6). Furthermore, Madeline was seen, to some extent, as a prototypical histrionic personality disorder ($M = 5.58$) and narcissistic personality disorder ($M = 5.47$), as was Charles for schizotypal personality disorder ($M = 5.01$) and Earnst for schizoid personality disorder ($M = 5.00$). The remaining three case vignettes failed to reach high degrees of prototypicality (mean ratings greater than 5.00) for any personality disorder.

FFM profiles. Table 4 presents the means and standard deviations for the FFM domains and all 30 facets for each case vignette. The ratings reported are based on a 7-point Likert scale. Ted scored fairly high on the Conscientiousness ($M = 4.87$) and Extraversion ($M = 4.77$) domains and low on the Agreeableness ($M = 1.70$) and Neuroticism ($M = 2.67$) domains. Madeline was high on Extraversion ($M = 5.89$) and moderately high on Conscientiousness ($M = 4.77$) and Openness ($M = 4.68$) but was fairly low on Agreeableness ($M = 2.49$). Earnst's only outstanding dimension on his profile was his low score on the Extraversion domain ($M = 2.12$). However, he did display moderately high scores on Conscientiousness ($M = 4.80$) and Neuroticism ($M = 4.76$). Samuel and Widiger (2006) provided mean ratings (using a 5-point Likert scale) for each facet but did not report mean scores for the domains. When averaging across the facets for the respective domains, Ted's highest ratings were on Conscientiousness ($M = 3.92$) and Extraversion ($M = 3.51$), and he scored the lowest on Agreeableness ($M = 1.45$) and Neuroticism ($M = 2.26$). For Madeline, her highest rated domains were Extraversion ($M = 4.19$), Conscientiousness ($M = 3.58$), and Openness ($M = 3.21$), with her lowest

ratings occurring on Agreeableness ($M = 1.78$). Earnst showed low scores on the Extraversion domain ($M = 1.69$), with his highest ratings on Conscientiousness ($M = 3.75$) and Neuroticism ($M = 3.68$). Overall, results for the FFM reported by Samuel and Widiger followed the same trends of the current study for these three case vignettes; however, Samuel and Widiger reported more extreme facet ratings.

Charles tended to show much fluctuation within the domains rather than between the domains. However, he consistently was described as being fairly high on Openness and low on Agreeableness and Conscientiousness. Aileen portrayed an obvious profile of low scores on Agreeableness and Conscientiousness, and Meursault showed a slight trend for fairly low ratings on all of the domains, especially Extraversion and Openness.

Marianne's FFM profile was mostly neutral, except for low ratings made in the Extraversion domain and slightly high ratings for Neuroticism. Alex showed high ratings for Extraversion, Openness, and Neuroticism, and moderately low ratings for Agreeableness. Finally, the psychologists provided Macon with low ratings for Extraversion and Openness and moderately high ratings for Conscientiousness. Overall, the case vignettes seemed to capture the full spectrum of all the domains, except for Agreeableness where a majority of the cases tended to receive low ratings, suggesting a general characterization of mistrust, manipulation, deception, exploitation, and ruthlessness across all the case vignettes.

GAF profiles. Psychologists generated three GAF scores for each case vignette—the most accurate GAF rating, the highest possible GAF rating, and the lowest possible GAF rating. Table 5 provides the means, standard deviations, and ranges for each type of GAF score for all case vignettes. Ted, Charles, Aileen, and Alex were described as the

lowest functioning with mean GAF scores less than 30. Madeline, Macon, and Earnst were considered the highest functioning, relative to the other case vignettes, with mean GAF ratings above 50. Using standard deviations as a measure of agreement, the psychologists were best with assigning GAF ratings to Earnst and Macon (*SDs* less than 10); whereas Meursault, Alex, Ted, and Aileen revealed the highest deviations among the psychologists (*SDs* greater than 15).

Reliability Measure for the DSM and FFM Ratings

Table 6 displays Cronbach's alpha for each case vignette across the *DSM* and *FFM*. This reliability measure was calculated using the raters as variables and the *DSM* categories or the *FFM* facets as cases. Overall, reliability was excellent for the *DSM* (α 's > 0.94) and the *FFM* (α 's > 0.93) across all case vignettes. Results were also comparable to Samuel and Widiger (2006) for Ted, Madeline, and Earnst, as the authors reported alphas greater than 0.95 for all three cases.

Clinical Utility Ratings

For each classification system, psychologists determined how confident they were in their generated profiles and how useful the respective system was when generating the profiles. These questions were assessed using a 7-point Likert scale (1 = not at all; 7 = very). Table 7 reports the means and standard deviations for the utility ratings. A one-way, repeated measures ANOVA found a significant difference among the three systems across all case vignettes in regards to the psychologists' confidence in their ratings, $F(2, 624) = 55.25, p < .001$. Pairwise comparisons revealed that psychologists were more confident in their *DSM* profiles than the *FFM*, $t(315) = 11.65, p < .001$, and the *GAF*,

$t(312) = 4.78, p < .001$. Furthermore, the psychologists reported more confidence with the GAF than the FFM, $t(316) = 5.33, p < .001$.

With respect to the usefulness of the systems in describing the patients, a significant difference was also found among the classification systems, $F(2, 626) = 52.01, p < .001$. The psychologists rated the *DSM* as being more useful than the GAF, $t(313) = 9.71, p < .001$. Further, the FFM was rated as more useful than the GAF, $t(317) = 7.14, p < .001$. No significant difference in usefulness was found between the *DSM* and FFM.

Sprock (2003) measured participants' confidence when using the *DSM* model versus the FFM and found similar results as reported here, concluding that participants were more confident using the categorical approach. Sprock measured other facets of clinical utility but these were not similar enough to compare against the usefulness measure utilized in this current study. In regards to usefulness of the systems, the results of the current study did not replicate findings from Samuel and Widiger (2006). These authors measured six aspects of clinical utility. Only two aspects appear to most closely resemble the usefulness measure employed in the current study: "How useful is this system for comprehensively describing all the important personality problems the individual has?" and "How useful was this system for describing the individual's global personality?" Samuel and Widiger reported significant differences in favor of the FFM for both these questions for all three cases. As stated before, the current study found no significant difference in usefulness between the *DSM* and FFM. Participant's confidence ratings of using the two systems were not measured by Samuel and Widiger.

Comparing Across the DSM, FFM, and GAF Profiles

A common metric system was needed in order to make meaningful comparisons across the three classification systems. Classification models represented in the same terms avail themselves to certain analyses to compare the variability in ratings among the classification systems and across the case vignettes. Variability values were calculated for each clinician across the three systems to compare, on average, how consistent psychologists were in making *DSM*, *FFM*, and *GAF* ratings. The variability values were determined by taking the absolute difference between the individual participant's rating and the sample mean rating for each case vignette using a particular classification system. Variability values were also examined for each system across all case vignettes to determine which system as a whole gave the most precise measurement. Such an analysis seemed the most appropriate, and meaningful, in terms of evaluating the actual distance psychologists' ratings fell from the mean rating for each classification system. The more this distance increased, then the greater the variability value, indicating that sufficient agreement did not occur for that particular system.

The 7-point Likert scale used for the *DSM* and *FFM* was the most obvious choice to generate comparisons, because the *GAF* can be rescaled from a 100-point scale to a 7-point scale (by multiplying each *GAF* score by 0.06 and then adding one). For example, a *GAF* score of 89 yields a new rating of 6.34, and a *GAF* score of 4 results in a rating of 1.24. Adding one ensures that the new *GAF* rating remains within the boundaries of the 1 to 7 point Likert scale (no scores will fall below a rating of 1). Therefore, variability was measured and compared among the psychologists and across the classification systems by using the 7-point Likert scale.

A variability value was calculated by taking the absolute difference between the mean prototypicality rating and the participant's prototypicality rating for the two most frequently selected *DSM* personality disorders across each case vignette. For example, Ted's two most frequent personality disorder diagnoses were antisocial and narcissistic. Ted's mean prototypicality rating for antisocial personality disorder was 6.29. The absolute difference was found by subtracting this mean rating from all individual psychologists' ratings of antisocial personality disorder for Ted, thus creating a variability value for each participant. The same method was used to find Ted's variability value for narcissistic personality disorder. A total *DSM* variability value for each case vignette was determined by calculating the mean absolute differences from the two most frequently diagnosed personality disorders. Column 2 of Table 8 displays the mean and standard deviation of the variability values using the prototypicality ratings for the two most frequently selected personality disorders across all case vignettes.

For the FFM scale, the mean rating of the six facets for each respective domain was calculated to determine five overall domain scores for each case vignette. The variability value was found by taking the absolute difference between each of the mean domain ratings and the participant's domain rating for each case vignette. For example, Marianne's mean domain rating for Neuroticism was 4.84 (determined by taking the average of the six facets comprising this domain). The absolute difference was found by subtracting this mean rating from all individual psychologists' ratings of Neuroticism for Marianne, thus creating a variability value for each participant. The same method was used to find Marianne's variability value for the remaining four factor domains. A total FFM variability value for each case vignette was determined by calculating the mean

absolute differences across the five domains. Column 3 of Table 8 presents the mean and standard deviation of the variability values for the FFM across all case vignettes.

As mentioned before, participants were instructed to provide the most accurate GAF score, as well as the lowest and highest possible GAF scores. Once the GAF scale was converted to a 7-point scale, the psychologists' lowest possible GAF scores were subtracted from their highest possible GAF score, resulting in a range value for each psychologist. For example, Charles' lowest possible average GAF score was 10.97 and his highest possible average GAF score was 34.06, thus resulting in a range value of 23.09. The variability value was found by taking the absolute difference between the total mean range GAF score and the participant's range GAF score for each case vignette. The fourth column of Table 8 depicts the mean and standard deviation of the variability values for the GAF across all case vignettes. The last column summarizes the total mean absolute differences for each case vignette across the three classification systems.

A one-way, repeated measures ANOVA found significant difference in the variability values among the three classification systems when including all case vignettes, $F(2, 632) = 206.91, p < .001$. Pairwise comparisons revealed that more agreement occurred among the psychologists when using the FFM scale, $t(316) = 17.05, p < .001$, and the GAF scale, $t(316) = 15.52, p < .001$, than with the *DSM*. Furthermore, more agreement occurred when using the GAF scale than with the FFM $t(319) = 2.25, p < .05$. In regards to specific case vignettes, significant differences in variability across the *DSM*, FFM, and GAF occurred across all case vignettes (all *F*-tests significant at $p < .001$, unless otherwise noted): Ted, $F(2, 70) = 7.04, p < .01$; Madeline, $F(2, 60) = 82.87$; Earnst, $F(2, 62) = 62.28$; Charles, $F(2, 68) = 22.73$; Aileen, $F(2, 80) = 16.60$; Meursault,

$F(2, 70) = 38.14$; Marianne, $F(2, 58) = 29.77$; Alex, $F(2, 70) = 3.88$, $p < .05$; and Macon, $F(2, 78) = 87.91$. For Charles, Aileen, Meursault, and Marianne, psychologists agreed more when using the FFM or GAF rather than the *DSM*. When describing Ted and Alex, psychologists performed better with the FFM than the *DSM*. In regards to Madeline, Earnst, and Macon, the best agreement was with the GAF scale, followed by the FFM, and then the *DSM*. Table 9 provides a summary of how each case vignette ranked among the three classification systems in terms of rater agreement.

When reviewing the mean total variability values across the three classification systems for each case vignette (refer to column 5 of Table 8), the psychologists had the most difficult time agreeing on the profiles for Meursault, Charles, and Marianne, as represented by the highest variability values. In regards to the *DSM*, psychologists struggled with agreeing on Meursault, Marianne, and Charles. The psychologists also did not reach high agreement with these same three case vignettes when using the FFM. For the GAF, the psychologists deviated from each other the most when generating GAF scores for Ted, Charles, and Alex.

Word-List Worksheet

Word descriptors generated by the participants were categorized based on whether the descriptor was found within the *DSM*, FFM, or GAF guides. This was achieved by using the “Find” function of Microsoft Word. For example, one participant described Meursault as “bizarre.” This word was entered into the “Find” search engine for the *DSM*, FFM, and GAF guides. In this case, “bizarre” appeared in the *DSM* and the FFM, but not the GAF; therefore, “bizarre” was coded as a *DSM* word and a FFM word. Due to the different word-list task procedures between Method I and Method II, analyses of the

word-list data will be described for the Alabama sample and then for the Idaho/Connecticut combined sample. Finally, the data then will be examined across all three samples. Table 10 provides a summary for each sample and the combined sample for the frequency of descriptors across all case vignettes for each classification system. Appendix B contains an example word-list of a participant from each state.

Alabama sample. A total of 1511 descriptors were provided by the Alabama psychologists. In Table 10, the totals for the three systems exceed 1511 because some descriptors fell in two or three categories. For example, Earnst was described as “rigid,” which appears in both the *DSM* and FFM lexicons, and therefore was coded twice. Furthermore, Marianne was described as “depressed,” which appears in all three classification systems, thus coded three times. When examining only pure *DSM*, FFM, or GAF descriptors (i.e., descriptors coded for only one classification system), the total number of words decreased to 573. Of these 573 words, just over 40% were *DSM* descriptors (e.g., narcissistic, obsessive, disturbed), almost 60% were FFM descriptors (e.g., methodical, cunning, reserved), and only one word was a GAF descriptor (able to hold down a job).

Idaho/Connecticut samples. As described previously, the Idaho and Connecticut psychologists provided descriptors at three different stages: the first word-list task, memory cues on a note card, and the second word-list task. Across these three stages, the psychologists provided a total of 1904 descriptors. When examining only pure *DSM*, FFM, or GAF descriptors, the total number of words decreased to 704. Of these 704 words, about 48% were *DSM* descriptors, just over half were FFM descriptors, and only about 1% was GAF descriptors. For the initial word-list task, the Idaho/Connecticut

participants' use of terminology is quite similar compared to the Alabama data, suggesting that the procedure order of the word-list task did not have an influence on the participants. Of the pure words used in the initial word-list task, 40% were *DSM* terms, 60% were FFM, and 1% was GAF. However, when examining the pure words for the memory cues, participants used *DSM* words two-thirds of the time versus less than a third for the FFM and about 3% for the GAF. On the follow-up word-list task, the psychologists listed about 51% pure *DSM* descriptors, just under half were pure FFM descriptors, and less than 1% was pure GAF descriptors.

Combined sample. When considering all three samples, a total of 3415 descriptors were provided by the psychologists. Of these 3415 descriptors, about 47% were coded as *DSM* descriptors, 49% were FFM descriptors, and almost 8% were GAF descriptors. When only examining pure descriptors, the participants listed 45% *DSM* terms, 55% FFM descriptors, and less than 1% was coded as pure GAF terms.

DISCUSSION

Unfortunately, the most striking piece of the current study is the low response rate. The study represented about 6% of all the psychologists recruited across the three states. This percentage rate is about three to four times lower as compared to studies similar in nature (25% from Samuel & Widiger, 2004; 19% from Samuel & Widiger, 2006; 19% from Sprock, 2003).

It is noteworthy to mention that these earlier studies used samples deliberately taken from organizations of practicing psychologists. For instance, the two Samuel and Widiger studies specifically recruited participants from the American Psychological Association Division 42, which is an organization of psychologists in independent practice. These authors were clear in their reason for recruiting from this sole division: "...to maximize the probability that the participants would be actively engaged in clinical practice" (2006, p. 300). Restricting the sample to psychologists fully involved in clinical work would likely increase the response rate of a study that examines clinical decision-making and utility. The recruitment processes of Sprock and Samuel and Widiger failed to include psychologists who are not entirely involved in traditional therapy activities, such as those psychologists in the settings of academics, forensics, medical/health, military, and corrective services. Conversely, the sampling procedure of the current study included psychologists practicing in a variety of venues and likely is a better representation of the psychology field as a whole. Though independent practice

psychologists may be more engaged on a daily basis with classification issues surrounding clinical decisions and utility, other settings of psychology still utilize and value the diagnostic process and should not be without a voice in this debate.

However, due to such differences in the recruitment processes, data gleaned from the samples are likely to have some discrepancies. When comparing results on Ted, Madeline, and Earnst between the current study and Samuel and Widiger (2006), the trends for *DSM* diagnoses and FFM descriptions were found to be quite similar. However, upon closer examination, the sample of Samuel and Widiger consistently rated these three case vignettes higher on almost all personality disorder diagnoses and FFM domains. This can clearly be seen by the percentages of *DSM* diagnoses for each case. To compare between prototypical ratings, Samuel and Widiger's ratings were converted to a 7-point Likert scale. Standard errors of the means were calculated and the differences between the mean ratings of the two studies were converted to *z*-scores. The majority of the *DSM* prototypical ratings for Samuel and Widiger were at least two standard deviations higher than ratings made by the sample of the current study. Therefore, Samuel and Widiger's participants rated the three case vignettes with the highest prototypicality rating (in the case of their study, a rating of 5) more frequently compared to the current study (a rating of 7). Such discrepancies occurred with just as much frequency for the FFM domains, where participants of Samuel and Widiger rated the cases with increased pathology.

There are likely two reasons for the occurrence of this finding. First, this finding could be a function of the rating scale. In regards to the *DSM* prototypicality rating scale, the participants of Samuel and Widiger (2006) may have been reluctant to provide a

rating of 4, as this rating was close to the middle of the continuum (a rating of 3), thus suggesting a “somewhat like” representation of the personality disorder. Therefore, it is feasible that participants relied more on a rating of 5 rather than a rating of 4 when making diagnostic ratings for those diagnoses that seemed to fit the case vignette. On the other hand, the participants of the current study had the opportunity to use a lower rating and still be able to represent the case as fairly high on prototypicality (a rating of 6). Therefore, this rating was used frequently, given that it was far from the middle of the continuum (a rating of 4) and it did not force the participants to consider the case as a pure prototypical case (a rating of 7). Consequently, participants of the current study consistently assigned lower ratings compared to Samuel and Widiger.

A second explanation may be the differences between the samples. As expressed before, Samuel and Widiger (2006) used a sample of practicing psychologists who are thoroughly involved in the diagnosing and assessment of patients. Therefore, this group may be more astute in recognizing the nuances of psychopathology, and thus are more apt to make more pathological ratings. The disconcerting fact, however, is that comparisons between these two studies may not be appropriate given the differences in how the two samples responded to the case vignettes.

A meaningful finding of such a low response rate is that it could be construed as an inadvertent measure of clinical utility. General psychologists who are likely involved in various activities may have found the demands of the study to be excessive. For those not familiar with the FFM (which was clearly seen by the familiarity ratings), having to learn the FFM and then apply it to several complex case vignettes could have made any well-intentioned psychologist discard such a time-consuming task. To empirically

measure this hypothesis, the psychologists who did not complete the materials could be contacted again and asked to provide feedback on the characteristics of the study and describe at what point they decided to not finish. Examining their reasons for not completing the study may prove to be helpful in identifying what was not clinically useful among the classification models.

Additionally, lack of incentive and the extensiveness of the study are likely contributors to the low response rate. Interestingly, when a personal cover letter was added to the second sampling procedure for the Alabama psychologists, the response rate of returned postcards doubled. This is consistent with findings from Yammarino, Skinner, and Childers (1991) who examined mail survey response rates. These authors performed a meta-analysis of studies using mail recruitment techniques and concluded:

“...researchers that use a cover letter that includes appeals...increase their response rates” (p. 627). However, these authors further found that response rates lower significantly when surveys are more than four pages and no incentives are offered. Given the extensiveness of the current study (as expressed by some unsolicited comments: “You have got to be kidding. I work 10-12 hrs. a day. I would never do this.” and “I now can’t take time for this extensive study.”) and the lack of compensation (“In the future, please offer some kind of an incentive for participation, a lottery raffle, etc.”), it is not surprising psychologists were not motivated to participate in this dissertation project.

Comparisons Across the Classification Systems

An unexpected result occurred when examining the variability values among the three classifications across all the cases. The GAF proved to be the measure with the highest level of agreement. The reason for this finding might be a reflection of the

extreme severity of the case vignettes, thus eliminating the top third range (GAF scores > 65) of the GAF scale. In regards to GAF scores, the range of the cases was 18-65, with over a half of the cases falling below a GAF of 40. Therefore, agreement among the psychologists would inevitably increase as the range of possible GAF scores decreases. The inclusion of higher functioning case vignettes, thus broadening the scope of possible GAF scores, may eliminate this finding. Despite such accuracy in using the GAF, the participants rated both the *DSM* and FFM as significantly more useful when describing the patients compared to the GAF.

An interesting trend existed with the psychologists' variability values for the GAF. The sample of psychologists tended to agree more on GAF ratings for the higher functioning cases (i.e., Madeline, Earnst, and Macon). On the contrary, the more dysfunctional the case (i.e., Ted, Charles, and Alex), the greater the disagreement when assigning a GAF score. An inherent problem with the GAF system is the obscurity of determining a specific level of functioning of an individual across a wide range of behavior patterns. For the more dysfunctional cases, these patients, despite their psychopathology, had periods of time when they were successful and productive. Therefore, psychologists had the complex task of assigning a single GAF score for Ted, a serial murderer, who also worked on several successful political campaigns and was seen by his colleagues as responsible, dedicated, and hard-working. Consequently, the participants of this study were not confident in their GAF profiles as compared to the *DSM* and FFM, nor found the GAF scores to be useful in accurately capturing an individual's entire personality presentation. The low ratings of clinical utility suggest the participants were frustrated with the GAF in generating a single score to represent a wide

repertoire of an individual's behavior. With increased emphasis placed on clinical utility, the GAF falls short of being seen as a comprehensive and helpful classification system.

Overall, participants were able to use the *DSM* and FFM systems with excellent reliability, as measured by Cronbach's alpha. This finding has been consistent with Sprock (2003) and Samuel and Widiger (2006). However, this study utilized a second method to evaluate the participants' consistency in ratings. When considering absolute differences between individual participant ratings and overall mean ratings, a greater consensus was reached with the FFM, as opposed to the *DSM*. In other words, participants were closer in agreement when assigning FFM facet ratings for the case vignettes than when assigning *DSM* ratings. Even when examining cases deemed as "classic" or "prototype" examples of *DSM* personality disorders (i.e., Ted and Aileen as antisocial, Alex as borderline), the participants agreed more on the FFM ratings for these cases as compared to the *DSM*. Notably, participants rated each of these systems as equally useful. Therefore, the significant finding of this study is that the FFM is the more superior model in terms of consistency in ratings and is seen just as useful as the current classification system when describing personality psychopathology.

Alex: A Case Example

The comparison among these classification systems may best be understood through a specific analysis of one of the case vignettes. The case vignette Alex was taken from the 1987 movie, *Fatal Attraction* (Lyne, Jaffe, & Lansing). The release of this movie generated much controversy and discussion, especially from feminist groups who were outraged that a powerful businesswoman (Alex) was portrayed as a psychotic stalker obsessed with a married man. Many saw this movie as a social comment on the

lifestyle of the 1980s, where family men had affairs and wives resentfully stood by them. This movie was also of great interest to the mental health profession, as Alex was discussed as the prototypical case of borderline personality disorder. Glenn Close, who played Alex in the movie, even visited with psychiatrists to gain a better understanding of borderline personality disorder in order that she may realistically portray Alex's character. However, the mental health profession was displeased with the released ending of the movie, as the final scenes did not correspond with the nature and behaviors of a borderline personality. In order to satisfy and entertain audiences, *Fatal Attraction* was released with an overdramatic, Hollywood thriller conclusion of Alex attacking Dan's family with a knife. After a lengthy fight scene between Alex and Dan, Alex was ultimately shot to death by Dan's wife. The original ending involved Alex taking her own life and framing Dan for her death. This less gory and suspenseful conclusion disappointed audiences but was a more realistic portrayal of a borderline personality character.

Fatal Attraction was one of the top grossing films in 1987. It was assumed most of the participants would be able to correctly identify Alex. However, just over 47% provided a correct response in identifying Alex. Such a low correct response rate may be attributed to the fact that the case vignette concluded with the original movie ending, rather than the horror-like movie conclusion. However, the percentage of correct responses was comparable to Aileen (49%), who was portrayed in a more recent, but less popular movie than *Fatal Attraction*.

Given the popularity of this movie and the continued reference of her character as a classic borderline personality disorder, presumably the participants of this study would

exhibit high levels of agreement when using the *DSM* to describe Alex's case history. Across all the case vignettes, more agreement occurred with Alex when determining a *DSM* personality disorder diagnosis. Clearly, the psychologists saw Alex as fitting into the category of borderline personality disorder. Interestingly, even such a "clear-cut" example of borderline personality disorder does not excuse it from one of the inherent problems of the *DSM* personality disorders, which is the overlap of diagnostic criteria. The participants also placed Alex in six other personality disorder categories, thus suggesting that she may meet criteria for well over half of the personality disorders. This problem of comorbidity was seen for all the case vignettes, as the minimum number of personality diagnoses given was five. Comorbidity has been a well documented concern of the *DSM* personality disorders, including those cases considered to be prototypic (Blashfield, McElroy, Pfohl, & Blum, 1994).

As mentioned before, participants had a difficult time agreeing on assigning GAF scores to Alex. Standard deviations ranged from about 15 to 20 across the three forms of GAF scores (the correct GAF score, lowest possible GAF scores, and highest possible GAF score), suggesting that Alex's level of functioning potentially could extend across multiple 10-point increment GAF levels. Again, it is difficult to assign a single GAF score to describe behavior that ranges from professional success to suicidal stalking.

As opposed to the GAF, participants were able to use the FFM to adequately represent specific features of Alex's pathological presentation, such as high levels of angry hostility, impulsivity, and excitement-seeking behaviors. Once these facets were well represented, participants then had the opportunity to portray more adaptive features of Alex, such as relatively normal levels of competency and achievement striving

behaviors, which speak more to Alex's success in her publishing career. Participants also were able to make clear distinctions among the facets within the FFM domains. For example, Alex was seen as high to problematic high on all facets comprising Neuroticism, except for the facet of self-consciousness. The participants were able to identify many neurotic behaviors of Alex without losing sight that she may not necessarily score high on all facets of Neuroticism. With the self-consciousness facet, participants recognized that though Alex can be rather intense and dramatic in relationships, she also possesses a relaxed and effective interactive style that allows her to be successful in developing relationships. This is a useful function of the FFM such that psychologists are able to be more specific and precise when describing a personality presentation because multiple ratings can be made within the same domain. Conversely, with a *DSM* categorical label, specific information is lost in regards to the presence and intensity of the criteria for the personality disorder that Alex exemplifies.

When comparing variability ratings between the *DSM* and FFM, the participants performed better when using the FFM to capture Alex, as compared to the *DSM*. This trend occurred across all case vignettes. Psychologists agreed more on the FFM domain ratings as compared to ratings made for *DSM* diagnoses. One might be surprised by this finding, given that the participants were clearly more familiar with the *DSM* and the GAF versus the FFM, as well as more confident in their *DSM* and GAF ratings compared to the FFM. However, given the complexity and richness of the "real-world" case vignettes, the FFM generally permits greater subtlety of personality psychopathology than the *DSM* and GAF in terms of conceptualizing wide ranges of behaviors, thus portraying a more comprehensive clinical picture (Lynam, 2002; Stone, 2002).

Word-List Analysis

The data from the word-list task suggest the psychologists used more pure FFM descriptors compared to pure *DSM* descriptors. This can be seen with the Alabama sample and the first word-list task done by the Idaho/Connecticut sample. When first presented with the case vignettes, the psychologists conceptualized the cases in more FFM terms compared to *DSM* terms. Therefore, the participants may find the FFM as a more natural classification system to describe the cases, which may have translated to the high levels of agreement seen with the FFM ratings. The FFM developed its roots through lexical studies, which provided a rich population of terms to describe adaptive and maladaptive attributes of personality (Saucier & Goldberg, 1996). Therefore, it may be no surprise that participants are more apt to use FFM terminology given the “...support for ongoing efforts to reconceptualize personality pathology within the framework of ‘normal’ personality theory” (Sheets & Craighead, 2007, p. 86; Costa & Widiger, 2002). The FFM innately employs a more natural language to communicate meaningful personality characteristics, whether these traits are functional or dysfunctional; whereas, the *DSM* structure contains terms and descriptions that exclusively focus on the pathology of the individual’s personality.

The word-list data for the memory cues and the second word-list task completed by the Idaho/Connecticut sample runs contrary to the dominant use of FFM terms, as discussed above. Preference goes to the *DSM* when needing to resort to quick prompts about a case, and both the *DSM* and FFM were equally represented when describing the case vignettes from memory. For example, with Alex (refer to Appendix B for a subset of Alex’s word-lists), the initial presentation of her case to the participants may have elicited

more pure FFM terms than pure *DSM* terms to describe her behaviors. However, when the participants needed to generate a quick reference of her case by writing memory cues, participants seemed to rely more heavily on *DSM* terminology than FFM. This can be expected as participants are more familiar and confident with *DSM* terms (i.e., borderline, obsession) to serve as simple cues to represent an entire clinical presentation. The *DSM* is known for its strength in communication and simplicity in using one-word labels to describe many symptoms. This primarily stems from the long-standing tradition of the *DSM* as the standard classification structure. However, tradition should not trump the validity, coverage, and applicability properties of a classification system. Clark (1993) and Verheul (2005) contended first and foremost that a classification system should be valid and a useful framework for clinicians and researchers. After this is achieved, the ease of communication and the confidence in the system will soon follow as the mental health field becomes more familiar with utilizing the classification system.

Limitations and Future Directions

One of the more difficult problems to address when examining different classification systems is finding a common measuring stick to adequately compare the systems. Each of these three classification systems contains their own unique properties, making it impossible to measure them against each other without changing the essence of the model. For the current study, the *DSM* was compared to the FFM and GAF in a way that is not its standard use. Instead of maintaining its categorical essence, the *DSM* was converted to a dimensional scale of prototypicality in order for comparisons to be made against the other two dimensional models. Therefore, caution must be made when concluding which system appears to be more superior in regards to clinical use when in

fact the system is not being used in its proper form. Two caveats need to be made explicit regarding the conclusion that the participants agreed more when making FFM ratings. First, participants were forced to use the *DSM* in a way that it was not intended to be used. It makes little sense to make definitive conclusions about a system if the properties of the system have been changed. Second, the FFM variability ratings were based on the domains, which consisted of the means of the facets. Consequently, the FFM domain values were the means of the means, thus inevitably reducing the variability among the ratings. This issue may be addressed by comparing FFM facet ratings to *DSM* personality disorder criteria ratings. Additionally, comparisons of variability values could be made between FFM domain ratings (as determined by facet ratings) against *DSM* cluster ratings (as determined by personality disorder ratings). As research continues to compare classification models, more creative techniques need to be employed to maintain the original structure and use of the classification system. This study attempted to do this through the word-list task, as well as the matching task (discussed in Appendix C).

A second limitation can be seen with the coding process of the word-list task. Words were coded based on the manuals sent to the participants. The *DSM* manual contained the entire contents of the personality disorder descriptions as found in the *DSM-IV-TR* (APA, 2000). However, participants received a modified version of the GAF and the FFM manuals, and coding was based using these modified versions. Therefore, words may not have been coded as a GAF or FFM word when in fact these words do appear when considering the entire contents of the GAF or FFM. However, it is unlikely that even if the entire GAF description was used, it would make up enough ground to match the volume of *DSM* and FFM words. Furthermore, the structure of the GAF does

not lend itself to much word descriptions; hence, it is not expected to be a comprehensive means in communicating personality profiles.

Interestingly, even though the entire FFM manual was not employed, it still outperformed the *DSM*. It would be expected that this difference would only increase if the entire FFM manual contents were used during the coding procedure. However, more behavioral evidence of clinical utility would be helpful to further recognize the distinctions between different classification systems in terms of how they are used to conceptualize personality presentations. The current study has begun to establish the foundation for such studies to examine further the unique features of the classification systems and how they can improve clinical practice.

Data gleaned from this study will serve to generate a new research thread in examining the clinical utility of the three classification systems. This next step will entail a task in which psychologists will be required to consider predetermined diagnostic profiles (for each of the three classification systems) and correctly match them to a case vignette (see Appendix C).

Sheets and Craighead (2007) expressed that research has firmly established that the "...structure of personality pathology have failed to replicate the Axis II organization" (p. 86). Therefore, additional studies are not needed to examine the current personality disorder classification, but rather empirical efforts need to be focused on identifying other models of personality organizations that are proven to be valid and clinically useful. The future direction of research on personality disorders, especially in light of the upcoming *DSM-V*, should emphasize how clinicians use classification systems to better understand,

communicate, and treat personality psychopathology (Verheul, 2005; Widiger et al., 2005).

Conclusion

No matter the endeavor, the development of a classification system stimulates closer observation and scrutiny of the objects being classified. Whether it is a young boy intently examining each rock to organize his collection based on size, color, or type; or a chemist carefully studying an element for its atomic weight, electron configuration, or metallic character in order to find its place within the periodic table, the classification of psychiatric disorders is no exception. Research studies are abundant in examining the intricacies of the psychiatric classification system. As the mental health field anticipates the arrival of the *DSM-V*, much research and attention has focused on the endeavor of classifying mental disorders, especially the personality disorders (Widiger & Samuel, 2005; Widiger et al., 2005)

The purpose of this research project was to assess how samples of psychologists could utilize different classification systems to describe case vignettes depicting personality disorders. This study is a modified extension of Samuel and Widiger (2006), with three important additions: (a) more case vignettes covering a wider range of personality pathology, (b) the inclusion of a third classification system (GAF) to act as the control, and (c) behavioral procedures to assess clinical utility, rather than sole reliance on self-report data. The goal of this research was not only to examine which classification system the participants utilized with the highest degree of consistency, but also which system seemed to make the most sense to the psychologists as they actively used them to describe the patients depicted in case vignettes.

The results of this study suggest that psychologists can employ the FFM system just as well, if not better, than the *DSM* system when describing personality psychopathology. Not surprising, participants were more confident using the *DSM*, due to the familiarity with the system, but participants equally viewed the *DSM* and FFM as useful models. Furthermore, participants utilized more FFM terminology when communicating information about the patients, likely due to the historical basis of the FFM in lexical studies. A significant strength of this study was the use of behavioral procedures to measure clinical utility, rather than relying solely on the participants' perspectives. With the arrival of *DSM-V* and the potential shift to a dimensional approach for personality disorders, more research is needed to examine how mental health professionals use different models of personality classification to improve communication, conceptualization, comprehensiveness, and treatment planning.

Future studies do not need to examine how clinicians perceive a certain classification model aides in communication, but rather empirical studies need to evaluate how clinicians actually use the model to communicate. Furthermore, future studies do not need to measure clinicians' perceptions of how a classification system assists in treatment planning, but rather how clinicians actually develop a treatment plan utilizing different classification systems. It is not sufficient to obtain clinical utility data through self-report methods, but rather assessing what clinicians actually *do* with these measurement systems is a richer test of clinical utility. The focus of such research will ultimately benefit the endeavor of improving upon the current classification system of personality disorders.

REFERENCES

- Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, 47(1, Whole No. 211).
- American Psychiatric Association. (1980). *The Diagnostic and Statistical Manual of Mental Disorders* (3rd ed.). Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (1987). *The Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., revised). Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (1994). *The Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (2000). *The Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text revision). Washington, DC: American Psychiatric Association.
- Ball, S. A. (2001). Reconceptualizing personality disorder categories using personality trait dimensions: Introduction to special section. *Journal of Personality*, 69, 147-153.
- Bernstein, D. P., Iscan, C., Maser, J., & Boards of Directors of the Association for Research in Personality Disorders and the International Society for the Study of Personality Disorders. (2007). Opinions of personality disorder experts regarding

- the *DSM-IV* personality disorders classification system. *Journal of Personality Disorders, 21*, 536-551.
- Blais, M. A. (1997). Clinician ratings of the five-factor model of personality and the *DSM-IV* personality disorders. *Journal of Nervous and Mental Disease, 185*, 388-393.
- Blashfield, R. K., McElroy, R. A., Pfohl, B., & Blum, N. (1994). Comorbidity and the prototype model. *Clinical Psychology: Science and Practice, 1*, 96-99.
- Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *Journal of Abnormal and Social Psychology, 38*, 476-506.
- Cattell, R. B. (1945). The principal trait clusters for describing personality. *Psychological Bulletin, 42*, 129-161.
- Clark, L. A. (1993). Personality disorder diagnosis: Limitations of the five-factor model. *Psychological Inquiry, 4*, 100-104.
- Clark, L. A. (2007). Assessment and diagnosis of personality disorder: Perennial issues and an emerging reconceptualization. *Annual Review of Psychology, 58*, 277-257.
- Clark, L. A., Livesley, W. J., & Morey, L. (1997). Personality disorder assessment: The challenge of construct validity. *Journal of Personality Disorders, 11*, 205-231.
- Cloninger, C. R., & Svrakic, D. M. (1994). Differentiating normal and deviant personality by the seven-factor personality model. In S. Strack & M. Lorr (Eds.), *Differentiating normal and abnormal personality* (pp. 40-64). New York: Springer.

- Costa, P. T., & McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., & Widiger, T. (Eds.). (2002). *Personality disorders and the five-factor model of personality* (2nd ed.). Washington, DC: American Psychological Association.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, *41*, 417-440.
- Endicott, J., Spitzer, R. L., Fleiss, J. L., & Cohen, J. (1976). The Global Assessment Scale: A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry*, *33*, 766-771.
- Eysenck, H. J. (1947). *Dimensions of personality*. New York: Praeger.
- First, M. B., Pincus, H. A., Levine, J. B., Williams, J. B. W., Ustun, B., & Peele, R. (2004). Clinical utility as a criterion for revising psychiatric diagnoses. *American Journal of Psychiatry*, *161*, 946-954.
- Galton, F. (1884). Measurement of character. *Fortnightly Review*, *36*, 179-185.
- Goldberg, L. R. (1972). Some recent trends in personality assessment. *Journal of Personality Assessment*, *36*, 547-560.
- Goldberg, L. R. (1976). Language and personality: Toward a taxonomy of trait-descriptive terms. *Istanbul Studies in Experimental Psychology*, *12*, 1-23.

- Goldberg, L. R. (1982). From ace to zombie: Some explorations in the language of personality. In C. D. Spielberger & J. N. Butcher (Eds.), *Advances in personality assessment* (Vol. 1, pp. 203-234). Hillsdale, NJ: Erlbaum.
- Goldberg, L. R. (1990). An alternative "description of personality": The big-five factor structure. *Journal of Personality and Social Psychology*, *59*, 1216-1229.
- Goldberg, L. R. (1992). The development of markers of the big-five factor structure. *Psychological Assessment*, *4*, 26-42.
- Goldman, H. H., Skodol, A. E., & Lave, T. R. (1992). Revising axis V for *DSM-IV*: A review of measures of social functioning. *American Journal of Psychiatry*, *149*, 1148-1156.
- Guilford, J. P. (1975). Factors and factors of personality. *Psychological Bulletin*, *82*, 802-814.
- Hay, P., Katsikitis, M., Begg, J., Da Costa, J., & Blumenfeld, N. (2003). A two-year follow-up study and prospective evaluation of the *DSM-IV* axis V. *Psychiatric Services*, *54*, 1028-1030.
- Hilsenroth, M. J., Ackerman, S. J., Blagys, M. D., Baumann, B. D., Baity, M. R., Smith, S. R., et al. (2000). Reliability and validity of *DSM-IV* axis V. *American Journal of Psychiatry*, *157*, 1858-1863.
- Kraemer, H. C., Noda, A., & O'Hara, R. (2004). Categorical versus dimensional approaches to diagnosis: Methodological challenges. *Journal of Psychiatric Research*, *38*, 17-25.

- Livesley, W. J. (Ed.). (2005). A research agenda for the development of a dimensional classification of personality disorder, part I [Special issue]. *Journal of Personality Disorders, 19*(2).
- Luborsky, L. (1962). Clinicians' judgments of mental health. *Archives of General Psychiatry, 7*, 407-417.
- Lynam, D. R. (2002). Psychopathy from the perspective of the five-factor model of personality. In P. T. Costa & T. A. Widiger (Eds.), *Personality disorders and the five-factor model of personality* (2nd ed., pp. 325-348). Washington, DC: American Psychological Association.
- Lynam, D. R., & Widiger, T. A. (2001). Using the five-factor model to represent the *DSM-IV* personality disorders: An expert consensus approach. *Journal of Abnormal Psychology, 110*, 401-412.
- Lyne, A. (Director), Jaffe, S. R. (Producer), & Lansing, S. (Producer). (1987). *Fatal attraction* [Motion picture]. United States: Paramount Pictures.
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*, 81-90.
- McCrae, R. R., & Costa, P. T. (1990). *Personality in adulthood*. New York: Guilford Press.
- Miller, T. R. (1991). The psychotherapeutic utility of the five-factor model of personality: A clinician's experience. *Journal of Personality Assessment, 57*, 415-433.

- Moos, R. H., McCoy, L., & Moos, B. S. (2000). Global Assessment of Functioning (GAF) ratings: Determinants and role as predictors of one-year treatment outcomes. *Journal of Clinical Psychology, 56*, 449-461.
- Murray, H. A. (1938). *Explorations in personality*. New York: Plenum Press.
- Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology, 66*, 574-583.
- Norman, W. T. (1967). *2800 personality trait descriptors: Normative operating characteristics for a university population*. Ann Arbor: University of Michigan, Department of Psychology.
- Samuel, D. B., & Widiger, T. A. (2004). Clinicians' personality descriptions of prototypic personality disorders. *Journal of Personality Disorders, 18*, 286-308.
- Samuel, D. B., & Widiger, T. A. (2006). Clinicians' judgments of clinical utility: A comparison of the *DSM-IV* and five factor models. *Journal of Abnormal Psychology, 115*, 298-308.
- Sanderson, C., & Clarkin, J. F. (2002). Further use of the NEO-PI-R personality dimensions in differential treatment planning. In P. T. Costa & T. A. Widiger (Eds.), *Personality disorders and the five-factor model of personality* (2nd ed., pp. 351-376). Washington, DC: American Psychological Association.
- Saucier, G., & Goldberg, L. R. (1996). Evidence for the big five in analyses of familiar English personality adjectives. *European Journal of Personality, 10*, 61-77.

- Saulsman, L. M., & Page, A. C. (2004). The five-factor model and personality disorder empirical literature: A meta-analytic review. *Clinical Psychology Review, 23*, 1055-1085.
- Sheets, E., & Craighead, W. E. (2007). Toward an empirically based classification of personality pathology. *Clinical Psychology: Science and Practice, 14*, 77-93.
- Siever, L. J., & Davis, K. L. (1991). A psychobiological perspective on the personality disorders. *American Journal of Psychiatry, 148*, 1647-1658.
- Skodol, A. E., Link, B. G., Shrout, P. E., & Horwath, E. (1988). The revision of axis V in *DSM-III-R*: Should symptoms have been included? *American Journal of Psychiatry, 145*, 825-829.
- Skodol, A. E., Oldham, J. M., Bender, D. S., Dyck, I. R., Stout, R. L., Morey, L. C., et al. (2005). Dimensional representations of *DSM-IV* personality disorders: Relationships to functional impairment. *American Journal of Psychiatry, 162*, 1919-1925.
- Sprock, J. (2002). A comparative study of the dimensions and facets of the five-factor model in the diagnosis of cases of personality disorder. *Journal of Personality Disorders, 16*, 402-423.
- Sprock, J. (2003). Dimensional versus categorical classification of prototypic and nonprototypic cases of personality disorder. *Journal of Clinical Psychology, 59*, 991-1014.
- Stone, M. H. (2002). Treatment of personality disorders from the perspective of the five-factor model. In P. T. Costa & T. A. Widiger (Eds.), *Personality disorders and*

- the five-factor model of personality* (2nd ed., pp. 405-430). Washington, DC: American Psychological Association.
- Thurstone, L. L. (1934). The vectors of mind. *Psychological Review*, *41*, 1-32.
- Trull, T. J. (1992). *DSM-III-R* personality disorders and the five-factor model of personality: An empirical comparison. *Journal of Abnormal Psychology*, *101*, 553-560.
- Trull, T. J., & Durrett, C. A. (2005). Categorical and dimensional models of personality disorders. *Annual Review of Clinical Psychology*, *1*, 355-380.
- Verheul, R. (2005). Clinical utility of dimensional models for personality pathology. *Journal of Personality Disorders*, *19*, 283-302.
- Watson, D. (Ed.). (2005). Toward a dimensionally based taxonomy of psychopathology [Special section]. *Journal of Abnormal Psychology*, *114*, 491-569.
- Widiger, T. A. (1992). Categorical versus dimensional classification: Implications from and for research. *Journal of Personality Disorders*, *6*, 287-300.
- Widiger, T. A. (1993). The *DSM-III-R* categorical personality disorder diagnoses: A critique and an alternative. *Psychological Inquiry*, *4*, 75-90.
- Widiger, T. A. (2000). Personality disorders in the 21st century. *Journal of Personality Disorders*, *14*, 3-16.
- Widiger, T. A., & Francis, A. (2002). Toward a dimensional model for the personality disorders. In P. T. Costa & T. A. Widiger (Eds.), *Personality disorders and the five-factor model of personality* (2nd ed., pp. 23-44). Washington, DC: American Psychological Association.

- Widiger, T. A., & Samuel, D. B. (2005). Diagnostic categories or dimensions? A question for the *Diagnostic and Statistical Manual of Mental Disorders—Fifth Edition*. *Journal of Abnormal Psychology, 114*, 494-504.
- Widiger, T. A., Simonsen, E., Krueger, R., Livesley, J., & Verheul, R. (2005). Personality disorder research agenda for the *DSM-V*. *Journal of Personality Disorders, 19*, 315-338.
- Wiggins, J. S. (2003). *Paradigms of personality assessment*. New York: Guilford Press.
- Wiggins, J. S., & Pincus, A. L. (1992). Personality: Structure and assessment. *Annual Review of Psychology, 43*, 473-504.
- Wiggins, J. S., & Pincus, A. L. (1994). Personality structure and the structure of personality disorders. In P. T. Costa & T. A. Widiger (Eds.), *Personality disorders and the five-factor model of personality* (pp. 73-93). Washington, DC: American Psychological Association.
- Yammarion, F. J., Skinner, S. J., & Childers, T. L. (1991). Understanding mail survey response behavior: A meta-analysis. *Public Opinion Quarterly, 55*, 613-639.

APPENDIX A: CASE VIGNETTES

Ted

Ted was born in Vermont to an unmarried mother. He never knew his biological father. He and his mother moved often in his first few years before settling in Washington where they lived with his uncle's family. His mother did eventually marry a man, but young Ted was apparently extremely jealous of this new target of his mother's affection. His mother and stepfather later had four more children, who he also apparently resented. Ted did well in school and typically earned A's in most of his classes, although he was sometimes in trouble for fighting with other children. He later attributed much of his scholastic success to the diligent efforts of his mother to encourage him. Despite his fondness for his mother, they never discussed personal matters and he stated that their relationship was not an open one.

Ted reported that he found it difficult to socialize and often chose to be alone or engage in solitary hobbies when in high school. Although, he was described as charming, intelligent and attractive, he had limited social contacts because he did not enjoy drinking and preferred the role of a scholar. He also had relatively few experiences with girls in high school and only went on one date. After graduation, Ted became involved with politics and worked on several successful campaigns, where he was described as being very responsible, dedicated and hard-working. Through this experience he was able to establish a wide social network. Ted used his charm and quick wit to establish himself as an up and coming politician and even was referred to as a "young JFK" for his political savvy.

At the age of 27, Ted began abducting, raping, and murdering young females. He often lured these women into his car by duplicitous means such as approaching them wearing his arm in a sling and asking for them to help him lift a boat onto his car or pretending that he was a police detective. The victims of these murders tended to be attractive university students. The acts of murder were quite brutal, including bludgeoning, mutilation and rape.

Ted's murders attracted a considerable amount of media and police attention, yet he continued to abduct women and evade detection. His success was attributed in part to his careful planning and execution of the abductions, as well as his charming and endearing manner with colleagues who could not imagine that Ted could ever be guilty of such acts. Ted was finally arrested approximately 15 months after his first murder. After his arrest, he was linked to several murders and was scheduled to stand trial. While in custody, his charm, good looks and cooperation soon won over his captors as they began

to provide him with special treatment, including the least restrictive restraints. He insisted upon defending himself in court and was allowed access to the local library where he diligently studied legal documents. Not surprisingly, on one such occasion when he was studying in the library he escaped by jumping from a second story window. He was later recaptured and kept under much more heavy guard. He proved to be a quick study in the field of law and was able to delay his hearings and trial for quite some time. During this delay he hatched a plan in which he was able to lose enough weight to fit his body through a 12-inch aperture and crawl through openings above the jail cells and offices.

After escaping, he soon settled near Florida State University. Not long after arriving he was once again raping, beating and killing several college women. During this time, he lived under a false name and was able to support himself by using stolen credit cards. He was eventually arrested after bludgeoning to death many of the members of a sorority house as they slept. He was subsequently found guilty and twice sentenced to death. However, his legal acumen was so high during the trial that after sentencing him, the judge stated that Ted “should have been a lawyer.” These legal skills continued to serve him while in prison as he delayed his execution for ten years. Others, however, suggested that his arrogant self-confidence contributed to a failure to obtain competent legal counsel and that his effort to serve as his own lawyer was ultimately harmful to his own legal defense. While in prison, Ted granted numerous interview requests and revealed that he committed the murders not as an act of violence, but instead as a means to having full possession of the women. In line with this, he claimed that the rapes were not brutal and that he had attempted to make the murders as painless as possible for the victims. Ted never expressed any explicit or compelling feelings of remorse for the murders. He, in fact, withheld the identities of many of his victims as a means toward delaying his execution.

Earnst

Earnst is a tall, slender young man of twenty-four. He was born to elderly parents as the youngest of nine children on a farm in rural Wisconsin. His father was a farmer by trade, but had taken various other manual labor jobs throughout his life to make ends meet. He described his mother as an intelligent, gentle, and loving woman, who was much thought of by friends and neighbors until her death when he was 14 years old. The family's fortunes were dealt another blow when his father became invalid soon after his mother's death when Earnst was 15.

His first recollections from a childhood he describes as largely miserable, were of being constantly sick. He reported that his father would at times drink heavily and become brutal towards him. The abuse he described was not physical in nature, but was verbally and emotionally disturbing. His father would curse and belittle Earnst, at times even to the extent of saying that his death would be a great benefit for the family. He recalls being quite sensitive to these criticisms and cried often after the verbal outbursts. Not surprisingly, he developed a feeling of inadequacy and often remarked that he believed something was wrong with him. On some occasions, these feelings of inadequacy still prevent him from looking other people in the eyes. Earnst also describes a tendency to perceive the laughter of strangers on the street as directed toward him. Although, after exploring the possibility further, he typically concludes that the impression was inaccurate. However, he states that these feelings seem to come up again and again.

Earnst possesses only a bitter hatred for the memories from his youth on the farm. With both his parents deceased he avoids contact with his many siblings. Further, he makes no effort to establish a new social network outside of this family unit, leaving him with a very solitary and lonely existence. Instead of pursuing relationships he focuses all his energies on achieving professional employment, which no one else of his family has ever attained.

As a child, Earnst first thought of himself as rather unintelligent, but soon found that he could excel scholastically relative to his peers. Upon realizing this, he devoted all his attention to his schoolwork and it "became the medium through which I was able to show my superiority and in a measure, to justify my existence." He was a student in the field of engineering, however he depleted his funds and was forced to postpone his education. During this time, he persistently sought a job in the field of engineering for which he was partially trained. Having spent most of his savings on education he was forced to take up a part-time job at a restaurant to earn his meals. Despite this depressive, hand-to-mouth existence, he insists on not accepting any job unless it is in engineering. Although, Earnst desperately wants to become a research engineer, he doubts whether he has the ability to be successful. Reflecting this uncertainty, he foresees only two possible outcomes of his life. He can envision himself as either achieving great success or becoming a dismal failure, with no gray area in between these extremes.

Earnst seems to be preoccupied with his lack of physical bulk. He describes instances as a child in which he was too puny to defend himself, even from children his own age, which caused him to develop a “fear of other people menacing me with physical punishment, which is something I have never entirely overcome.” His cowardice upset him so much that his life became intolerable, but because of this fear and his lack of size, he could do nothing to change it. He stated that he began to avoid the company of others as much as he possibly could. Instead, he found himself immersed in the lives of various heroes he would read about in books and was able to experience through them the recreation and enjoyment that was not coupled with the physical exertion of real activity. Because of this enjoyment of fantasy, he states that he never gets bored, but instead can “dig up something to do at almost any time.” However, in recent years he has experienced vision problems that have caused him to get away from reading quite so much and instead he attends theater shows for recreation.

Earnst states that he is “in general, retiring when with a group” and never assumes a leadership role or becomes the life of the party. He states that he never had a date with a girl until his third year of college. However, after this date he found that he enjoyed this feminine companionship and sought out the company of girls. Since that time he has spent much effort trying to figure out how to “arouse the interest of a girl,” but eventually found that relying on luck was his best option. Earnst describes eventually falling in love with a girl, “who happened to have some regard for me,” whom he became engaged to. However, soon after exhausting his savings, his fiancé left him for another man who was in better financial circumstances. He states that he has never had sexual intercourse and that he would not be interested in having sex with anyone he felt strongly for, unless they were already married to him.

Madeline

Madeline is a 35-year-old woman, currently employed as a defense attorney. Madeline grew up in a small, rural town. Both of her parents had serious drinking problems. She states that she and her siblings had to assume many of the adult responsibilities. Even worse, she describes her father as being severely physically abusive when intoxicated, toward her as well as her siblings and mother. She describes an instance as a child when she was asked not to wear shorts in gym class because the other children were frightened by her many bruises. She so strongly detested her father that she at times even harbored thoughts of killing him. Madeline explicitly recalled a time in which she had devised a plan to suffocate her father with a pillow, but decided not to carry it out upon realizing she was not physically powerful enough. Overall, she states that there was little “human warmth” in the family, from either her mother or her father.

She drank heavily herself as a child. By third grade she was taking whiskey to school in a thermos. She described herself as being wild and incorrigible, often running afoul of the law. She left home at age 12, alternately living in foster homes and jail. As she described it, “I did everything,” including theft, drug abuse, and fighting. She states that she had an epiphany while in jail as a young adult when she was left in solitary confinement after being hosed down following a jailhouse fight. She resolved to gain control of her life and never return to jail. This became something of a turning point as she began reading voraciously and educating herself.

At the age of 21, Madeline moved to another state and started over. She took classes at a community college and studied hard. She got married, but was soon divorced, as she indicates that the relationship was based mainly on “good sex.” Her good grades led to a scholarship at a major university where she eventually obtained a law degree, specializing in the defense of ethnic minorities and the disadvantaged. Her private practice grew rapidly and she soon established a very successful and well-regarded practice. She never made any effort to contact her parents or siblings, and expresses no interest in doing so.

She is described by her peers as being fun loving, outgoing, and at times even quite exhibitionistic. She is regularly flirtatious and appears to even enjoy the discomfort engendered in others by her often audacious and brazen behavior. She will “flash” men at parties (e.g., exposing her breasts) and make suggestively sexual remarks. Some persons find her entertaining, but others are quite offended. Madeline works and plays tirelessly. She seems to require little sleep, living life to its fullest, but also continues to build a strong and successful law practice. She desires not only to succeed but to be the very best at everything she does.

Madeline has a very active social life, hosting exciting and rowdy parties that would be the talk of the town. She reports pride in her variety of friends from all walks of life, including “strippers, musicians and college professors.” However, the quality and depth of her relationships are questionable. She would work tirelessly for a friend,

providing whatever instrumental assistance she could provide (which can be quite a bit), but she is not someone from whom one seeks emotional support or sympathy. She is quite up front in her willingness to do whatever it takes to win. She is described by others as a “master of manipulation,” who is capable of using any number of crafty and cunning maneuvers to get what she wants. She is unlikely to sacrifice her own desires to satisfy or mollify others. Madeline is also described as being boastful and self-aggrandizing. She does have good reasons to be proud of her accomplishments, but she is also known for grossly embellishing her background and history to increase the admiration of others as well as to impress upon others her current importance (e.g., telling stories about current friendships with famous musicians).

Her blazing success as a lawyer led eventually to an offer from a major law firm. At that point in her life she appeared to have no limits to her potential success. This new firm would offer a substantial increase in salary, recognition and status within the community, and an affirmation of her skills as a lawyer. However, she did not stay long with the prestigious firm. After only a few weeks, it was agreed that she “wasn’t capable of being an employee” (the details of the falling-out are not entirely clear).

As her private practice grew, Madeline fell in love with a man whom she described as her bedrock and foundation. She described only one serious prior relationship and it did not end well. She fully expected this relationship to be successful. They rented a house together and it appeared, at least to her, that this relationship would last forever. However, apparently, he was not as happy or satisfied with the relationship as she. He eventually bought a house surreptitiously, not informing her of his dissatisfaction or his intention of leaving. He did, however, tell her in person the night he left when she had returned from a business trip. Madeline was dismayed and baffled. They spoke a good deal about the relationship that evening. The one statement from him she remembers to this day is his exclamation that “it has to be about me now.” His departure was an emotional shock (one evening she sat staring at a bottle of sleeping pills), but she never felt she needed psychiatric treatment. She soon became absorbed again in her work and in her partying. Her practice continues to grow (“I keep winning and that keeps me happy”). She states that she will always love him and is not interested in finding a new relationship.

Charles

Charles was born in Illinois and was the youngest of three children. His mother died when Charles was seven years old, and he was raised by his older sister and father. Charles' father was a very successful businessman and flourished as a banker and county clerk, thus being a very neglectful and distant father. He also was a very harsh religious man who viewed all behaviors as either coming from God or the Evil One. Charles received severe punishments from his father when his behaviors were determined to come from the Evil One. Charles had a troubled childhood, with no rescue from his abusive father.

Charles decided to leave home and pursue his education at a college in Michigan when he was 18 years old. However, he quickly became very lonely and discovered that he was not prepared for university work. Charles found consolidation in religion and soon became interested in a religious community in New York. Just after a few months residing in Michigan, Charles left for New York to join the Oneida Community. Charles had a difficult time fitting in with the Oneida Community. He frequently lost his temper with many of the members and spent most of his time isolated, writing furiously and then burning what he had written. Many considered him incompetent and unreliable. At the age of 24, Charles left the community when it was revealed to him by God that he was "chosen" to be the editor of a daily, theological newspaper. Charles claimed that he was inspired and was "in the employ of Jesus Christ and Company."

After three months away from the Oneida Community, Charles reapplied for admission, stating that he wanted to become a "little child" and devote himself to the Community spirit. Charles expressed his loyalty to the Community by donating his full inheritance to the Oneida printing department. However, Charles once again left the Community after only about a year and demanded his inheritance back, which was refused by the Community. He left for Chicago to live with his sister and brother-in-law, but after only a few months Charles returned to New York to continue his pursuit in journalism. Unsuccessful in New York, Charles moved back to Chicago, was admitted to the Illinois bar, set up a law practice, and then married when he was 28 years old. Charles was an inadequate provider and companion, as he avoided paying bills and often abused his wife, striking her, pulling her by her hair, and locking her in the closet all night. A couple of years later, Charles and his bride moved to New York, though they divorced after five years of marriage. Charles purposely slept with a prostitute in order for his wife to have legal grounds for a New York divorce.

Once in New York, Charles became consumed in the presidential campaign, and he began writing campaign speeches and attending public meetings, convinced he would be appointed as the minister to Chile. However, Charles' political services were never called upon. Abandoning his political aspirations in New York, Charles left for Chicago to direct all of his attention to theology. He became convinced that he was chosen to preach a new gospel, and Charles began moving from town to town speaking at revivals and spreading his newly proclaimed message. He continued to leave a trail of unpaid

bills and his preaching primarily consisted of him shuffling papers, speaking incoherently for a few minutes, and then rushing out of the room. He even wrote a religious book that was eventually discovered to be almost entirely plagiarized.

Charles' evangelical mission continued for about four years until he became fascinated again with politics. He moved to New York and became a passionate follower of the Republican Party and its presidential candidate. Charles frequently confronted prominent politicians offering them campaign speeches he had written. When the Republican candidate was elected, Charles' bombarded the President and his staff with letters requesting position appointments, such as being the minister to Vienna. Charles soon moved to Washington, D.C. and frequented the State Department and White House, submitting numerous applications to become the minister of a handful of countries. He was certain that the President owed his successful election to him.

Charles continued his haphazard way of life in Washington, avoiding bill collectors and not having a source of income or any friends. Charles soon found himself banned from the White House, as his antics were no longer amusing to the presidential staff. At this point, it occurred to Charles that the President should be removed from his office, and that he himself was the chosen one for the job. Charles considered it a "divine pressure" for him to be the President. Charles purchased a cheap, wooden-handle revolver; however, he envisioned using a .44 caliber collector's revolver as it would be a more suitable museum exhibit after the assassination. On an early Saturday morning, when the President emerged through a reception room at the train station, Charles jumped out and fired two shots at the President, crying "I am a Stalwart of the Stalwarts! I did it and I want to be arrested! Arthur is President now!" The President died 80 days later from the gunshot wounds.

Charles served as his own defender when his case went to trial, and his court strategies were just as tangled and tragic as his life had been. Most of his testimonies were in the form of epic poems, and he typically sought legal advice from random audience members in the courtroom by passing them notes. Charles was found guilty and before he was to be executed, Charles read aloud some verses he had written, indicating that "if set to music they may be rendered very effective." He proclaimed, "I am going to the Lordy, I am so glad, I am going to the Lordy. Glory hallelujah! Glory hallelujah! I saved my party and my land. Glory hallelujah! But they have murdered me for it, and that is the reason I am going to the Lordy. Glory hallelujah! Glory hallelujah! I am going to the Lordy!"

Alex

Alex was a 36-year-old, attractive female who is an editor of a publishing company. She was the oldest daughter of a wealthy socialite woman who was known in society circles for her bad luck with men. Alex's mother was an extroverted woman who was married five times and died while living with a sixth man. The man who was most like a father to Alex was her mother's third husband. As a child, Alex was an introverted, dependent, and awkward girl. This step-father was the only person to show her attention and affection. Alex's mother and this man stayed married for 10 years, a record for her mother. However, the step-father started having intercourse with Alex when she entered puberty. Alex told her mother who refused to believe Alex. However, when the man had an affair with the mother's attractive personal secretary, the mother filed for divorce.

Alex was an intelligent woman with a good sense for business. Alex went to college at Swarthmore and then obtained her MBA at the University of California at Berkeley. After college, Alex moved to Malibu, California where she obtained her first job working for a firm that handled the careers of movie stars. While in Malibu, Alex became fascinated with jazz and learned about a late night DJ who played jazz on a local radio station. Alex started calling the DJ and asking him to play a song called Misty for her. Alex learned what bar the man frequented, and she approached him. Her intelligence, her charm, and her knowledge of jazz impressed this rough, self-confident man. They became lovers. After a couple of weeks, the DJ tired of Alex and tried to brush her off. She became obsessed with him and started stalking him. She appeared at his house and threatened suicide if he did not take her in. They resumed their affair. He became more desperate in his attempts to end the relationship. She eventually did make a suicide attempt after trashing his house, and she was hospitalized.

After the hospitalization, Alex moved to New York where she was hired by a publishing company. There she met Dan who was an attractive, narcissistic attorney who worked for the same publishing company. They met at a company party where they talked only briefly, but they discreetly caught glimpses of each other across the room throughout the party. The following day, Dan and Alex saw each other at a business meeting, and they continued their flirting. They agreed to have dinner together, and during dinner Alex revealed that she stood up her original date for that night to spend time with him. Dan added to the sexual tension between the two by lighting her cigarette and revealing that his family was away for the weekend. The two went to her apartment where their sexual affair started.

They continued their intense affair throughout the weekend. Their sexual acts were very aggressive, uninhibited, and risky. On their final night together, the two had dinner. Alex asked Dan about his family. He disclosed that he has been married nine years and that he and his wife have a six-year-old daughter. Alex expressed her frustration that all the good men are always married, and she attempted to convince Dan to continue with their love affair. Dan indicated that their relationship cannot continue,

and Alex appeared to understand his plight. They spent one more night together for what was presumed to be the last time.

After their night together, Dan began to dress to leave. Alex became visibly upset, claiming that he always makes love to her then leaves. As he continued to button his dress shirt, Alex flew into a rage. Dan tried to calm Alex down and to explain why their affair must end, but she became more enraged and refused to hear his explanations. Once dressed, Dan said goodbye and began to walk out the door, but not before he heard an overly sweet and soft-spoken Alex telling him that she wants to be friends. She apologized for her actions. She cried as he hugged her goodbye. She apologized again for becoming so angry and she began to kiss him passionately, putting her hands all over his face and through his hair. At this moment, Dan realized that there was blood on her hands and that she had slit both of her wrists. Alex cried hysterically as Dan frantically washed off the blood. He bandaged her wrists and stayed the rest of the night to take care of her. As he left the next morning, Alex very calmly said goodbye to him, asking that he call her sometime.

When Alex did not hear from Dan, she began to harass him at work, calling him incessantly. Dan told Alex that she cannot continue this and that he did not want to talk to her again. Alex then began calling Dan's house, but hung up when his wife answered. At one point, she called at two in the morning, stating that she had no choice to call him at home since he refused to accept her calls at work. Desperate to keep Dan in her life, Alex told him that she was pregnant with his child, even daring Dan to call her gynecologist for proof. Dan offered to pay for an abortion, but Alex insisted on keeping the baby, hoping that Dan would become a part of her family. Alex instructed Dan that their relationship did not have to be a problem, stating that, "You play fair with me; I'll play fair with you."

Unable to convince Dan to leave his former life and start a new family with her, Alex conceived the ultimate revenge to ruin Dan's life. Aware that Dan's fingerprints were on her kitchen knife from their weekend dinners together, Alex sat on her bathroom floor and cut her throat with the knife. A smile came across her face as she pictured her diary lying on her bed detailing her relationship with Dan. In her journal entries, she portrayed him as a possessive and abusive lover, who frequently stated that if he could not have her than no one could. Alex was sure the police would find the evidence, and Dan would suffer the rest of his life from her suicide.

Marianne

Marianne is a 38-year-old, single, obese woman who works at a mortuary. She has been employed at the mortuary since the death of her mother. Her mother and father were divorced when Marianne was 15 years old. Marianne's father used to laugh at his wife's dreams and ambitions to explore the world. Marianne was always very close to her mother and disinterested in her father, as he was with her. Initially, the father won custody after the divorce, but when Marianne kept running away from him she was permitted to live with her mother. After some time, Marianne's mother began to suffer from cancer, and Marianne took care of her until her mother died when Marianne was 21 years old. At the funeral, Marianne's father told her that it was their lifestyle that caused the death of her mother.

Marianne requested a job with the mortuary that had buried her mother. She felt an obligation to take care of the dead, stating that, "I protect the corpses just as I protected my mother." Marianne's lifestyle is very uneventful and bland. She has little to no contact with other people, including the other employees at the mortuary. She mostly keeps to herself and rarely speaks to anyone. She lives as though others around her do not even exist. The only joy Marianne finds in life is eating her dinners in bed while watching TV. Her daily routine consists of riding the subway back and forth to work and falling asleep each night after her dinner, with the glow of the TV flickering throughout her apartment room.

However, on one occasion while riding on the subway, she becomes fascinated with the voice of one subway driver and is determined to learn about this man. That very night Marianne returns back to the subway all dressed up with a hint of make-up on. She rides the subway he is driving through all its stops, eventually moving closer to the front subway car as passengers ahead of her leave. Finally, she faces the driver whose voice has intrigued her, but she can only stare at him. He glances at her, but then gets back into the subway and drives off.

Marianne becomes obsessed with trying to find him again and begins to miss work. She requests vacation time, indicating that she has one and a half years built up but only needs eight weeks time off. Marianne buys a subway schedule and sets her watch to match the subway clocks. She also buys a stopwatch to time the arrivals and departures of the subways to be sure the schedule is accurate. Marianne goes through painstaking measures to obtain the driver's work schedule and routes. She breaks into the personnel office and steals the drivers' schedules. At home, Marianne repeatedly goes over all the schedules and routes, using a color-coded system to map out his exact schedule. During this time, she also prepares for their eventual meeting by buying a new, bigger mattress, perfume, and lingerie. Marianne begins to hang around the subway station and watches him from a distance, never able to approach him. She even follows him home one day and lingers outside his apartment door, where she discovers that he is married. She buys binoculars so that she can spy on him at his apartment and learn more about his and his wife's daily routine.

Finally, she plans a meeting with him at the subway station by a candy machine where she proceeds to buy him a candy bar. He accepts the candy and leaves, only to find her there the next day with another candy bar. However, during this second encounter Marianne invites the driver for coffee. By the end of their time together, Marianne gives him her address and tells him to come to her apartment the next night for dinner.

Marianne prepares an elaborate meal and she is dressed very provocatively, as she waits for him to come. As it gets later, Marianne becomes more and more agitated since he has not yet arrived. She paces frantically in her apartment, going back and forth from the balcony to the stairwell anticipating his arrival. Eventually she gives up and concedes that he is not coming. Marianne begins to cry and eat all the food she has prepared. She then becomes violently angry and begins to bang a clock on the dining table, eventually breaking a wine bottle and a light fixture. After her rant, the driver shows up at her door claiming that he had to substitute for another driver. He comes in to the apartment and Marianne begins to seduce him. She takes his and her clothes off and he spends the night with her.

Marianne and the driver continue their love affair while the driver's wife is out of town. Marianne makes dinner for him every night and completely spoils him, showering him with gifts and attention. They spend much time talking and disclosing personal events in their lives. Marianne learns from him that a major fear of subway drivers is that someone would leap in front of the subway in a suicidal act.

The driver's wife returns and discovers that her husband is having an affair. She finds the two lovers at a nightclub and confronts Marianne. She immediately begins beating Marianne on the dance floor in front of everyone. Marianne collapses to the floor, crying in pain. The driver stands there and watches, doing nothing as his wife kicks and hits Marianne. The driver and his wife leave the club, as Marianne stays curled up on the dance floor crying.

Distraught and broken that her relationship with the driver has ended, Marianne puts on her best dress, makes up her face, and goes to the subway station the following day. She buys a candy bar, stands near the subway track, and dangerously leans over the track holding out the candy bar, waiting for her subway driver to arrive.

Aileen

Aileen was born to a 16-year-old mother and to father who she never knew. Aileen's father was a psychopathic child molester who was in prison all through Aileen's life until he committed suicide when she was 13 years old. Aileen had an older brother of one year named Keith. Aileen's mother could no longer handle the responsibilities and demands of single motherhood, and she abandoned Aileen and Keith, who were then adopted by their maternal grandparents. Aileen was four years old at the time. The grandparents raised the children as though they were their own, never telling them who their biological parents were. Aileen and Keith lived a troubled and rebellious childhood life. When she was six years old, she severely burned and scarred her face while playing with lighter fluid and matches. Aileen's grandfather was an alcoholic and would viciously beat Aileen with a belt, making her pull down her shorts and bend over the kitchen table. Most nights she would lay face down, naked on her bed, wincing in pain from the beatings. Aileen and Keith both discovered the truth about their grandparents not being their biological parents when Aileen was 12 years old, resulting in an even more rebellious attitude.

Aileen was sexually active at a very young age; some speculate her first sexual experience was with her brother, Keith. She became pregnant at age 14 and was sent to an unwed mothers' home. She delivered a boy who was taken from her and given up for adoption. Aileen never saw her son again. In that same year, Aileen's grandmother died, supposedly of liver failure. Apparently, the grandmother started having violent convulsions at home, but the grandfather refused to call the ambulance as he did not have the money to pay for it.

Aileen decided to drop out of school, left home, and took up hitchhiking and prostitution. In the next few years, Keith died of throat cancer when he was 21 years old and Aileen's grandfather committed suicide. Aileen was 20 years old and headed to Florida to make money in prostitution. She met a wealthy 69-year-old yacht owner, who quickly fell in love with her. They married, but Aileen was too wild and destructive, treating her husband badly, getting into bar fights, and was sent to jail for assault. After a month of marriage, Aileen's husband had the marriage annulled.

For the next ten years, Aileen staggered from one failed relationship to another. She remained active in prostitution, as well as forgery, theft, and armed robbery. She even at one point attempted suicide. Distraught over a failed relationship, Aileen got drunk and bought a gun with the intent of shooting herself. However, she lacked the courage to do so and ended up robbing a supermarket dressed in a bikini. She served 18 months in prison.

Aileen continued her work in prostitution and on one occasion one of her customers got too aggressive with her and started to beat her and sodomize her. In desperation, Aileen found a gun in the man's car and fired three shots at him, killing him

instantly. Aileen found killing to be a great way to express her anger towards men and the world, and she began to enjoy the control she had over her victims.

Over the next several months, two more bodies were found. In all, Aileen killed seven men along the Florida highways in less than a year, with three killings occurring within a three week period. Predator-like, she flagged down strangers while hitchhiking and once in their cars would offer sex. However, before any sexual acts occurred, Aileen would shoot them several times, and then she robbed them and left their naked bodies hidden in the woods.

Aileen was arrested at the age of 35. During her confession, Aileen reiterated that none of the murders were her fault and that all the killings were done in self-defense. She described that all of her victims assaulted her, threatened her, and raped her. Aileen explained that whenever one of her victims became aggressive with her, she would shoot him out of fear. A media frenzy soon erupted and book and movie offers were pouring in to Aileen. She believed she was famous and found herself all over the local and national news. She continued to talk about the crimes with anyone who would listen, including the jail employees. Each retelling of the story was elaborated upon and glorified, casting her in a better light each time.

Aileen went to trial, and if convicted she would receive the death penalty. The jury was made aware of her killing spree and her videotaped confession. The jury noticed that on the videotape Aileen appeared confident and not at all upset when recounting the murders. She easily conversed with the interrogators and frequently told her public defender to shut up. She proclaimed at one point to the video camera, "I took a life. I am willing to give up my life because I killed people. I deserve to die."

In less than two hours of deliberation, the jury reached a verdict. They found Aileen guilty of first-degree murder, to which Aileen exploded in rage, shouting, "I'm innocent! I was raped! I hope you get raped! Scumbags of America! I hope your wife and children get raped in the ass!" Aileen was sentenced to die in the electric chair.

Aileen did not appeal the court decision, indicating that she wanted to get right with God. She ultimately confessed that her first victim did beat and rape her, but that her other victims did not. She fired her attorneys and wrote a letter to the Florida Supreme Court stating, "I'm one who seriously hates human life and would kill again." Aileen requested to die by lethal injection instead of the electric chair, to which her request was granted. Aileen even demanded an earlier execution date. She died at the age of 42, her last words being, "I'd just like to say I'm sailing with the Rock and I'll be back like Independence Day with Jesus, June 6, like the movie, big mothership and all. I'll be back."

Meursault

Meursault, a single man in his early 30s, was just notified that his mother died at the nursing home she was residing in. The message was not clear on the exact date in which she died, but Meursault found that to be an inconsequential detail. His first thoughts were that he would have to ask off for work and that he would have to try and explain to his boss it was not his fault he had to take time off.

Meursault arrived at the nursing home. He made very infrequent trips to visit his mother, as it was a hassle to buy a bus ticket and endure the long drive. It would take up his Sundays, the day he enjoyed resting. It was not that he did not like his mother; she actually lived with him for a few months before he transferred her to a nursing home. But their lives were so different. She was old; he was young. She was a hindrance. He felt it was better for her to be around people her own age. At the mortuary, housed within the nursing home, the caretaker offered to open the casket so that Meursault could see his mother. Meursault refused. He sat with the caretaker and drank coffee and smoked a cigarette in the room where his mother laid.

The caretaker began setting up chairs for the other nursing home residents to pay their respects during the all-night vigil. Meursault dozed off while sitting and waiting, and he was roused as the residents began to enter the room. Meursault quietly watched the others file in, observing how fat the women were and how skinny the men were. He watched as some began to cry, wishing he did not have to listen to their sounds anymore. When the crying eventually stopped, Meursault found himself even more annoyed with the silence and the staring from the others at him, as though they were expecting him to display some grief or sorrow. Instead, Meursault's thoughts were preoccupied with how the silence magnified the coughing and lip smacking of the residents. Meursault sat quietly throughout the vigil, not speaking to anyone, occasionally dozing off.

The next morning the funeral director asked Meursault if he wanted to see his mother one last time before the casket was permanently sealed. Meursault refused again. Throughout the funeral procession, Meursault's attention was absorbed by the sweltering heat of the day. His agony and distress were a result of the sun unmercifully beating down on him while walking to the funeral site. Relief only came to Meursault when the funeral ended and he knew he could go inside and sleep.

The next day back at home, Meursault went swimming at the beach. He encountered a former co-worker that he always had a crush on. Her name was Marie. They decided to see a comedy movie that night, and Marie ended up spending the night with Meursault. The next morning Marie was gone by the time Meursault woke up, and he spent the day secluded in his apartment, watching the city activity from his balcony. He avoided making any trips outside his apartment as he did not want to answer questions from others about his mother's death. He concluded that nothing really had changed in his life, as he looked around his bland apartment. His mother was dead, but he still had to wake up the next morning and go to work.

Meursault and Marie continued seeing each other throughout the week, spending most of their time at the beach or at Meursault's apartment. At one point, Marie asked Meursault if he loved her, to which he replied, "It didn't mean anything, but I don't think so." Marie was startled by his response but laughed it off. Next door a fight had broken out between Meursault's neighbor, Raymond and his mistress. The mistress screamed out loud enough for the other residents to come out and peer down the hallway, including Meursault and Marie. Marie commented on how horrible the situation was; Meursault was not too concerned. Marie pleaded for Meursault to call the police, but he indicated that he did not like cops. However, a cop eventually showed and took Raymond to the police station. Marie was very upset by what just unfolded and was unable to eat her dinner. Meursault, on the other hand, cleaned his plate.

The next day Meursault was confronted by Raymond and invited Meursault into his apartment. Raymond asked if Meursault would act as a witness and testify that his mistress was cheating on him, thus the reason for beating her. Meursault agreed to do so, thinking, "Why not?" Raymond invited Meursault to his beach house for the weekend, stating that he could bring Marie. Meursault agreed to go, looking forward to more swimming and basking in the sun with Marie. At the beach house, during one of their long walks on the beach, Marie asked Meursault to marry her. Meursault replied that he did not care if they got married, and he emphasized again that he did not love her. He continued explaining that he would marry another woman in the same situation; it just did not matter to him.

One afternoon on the beach, two men approached Meursault and Raymond, one of the men being the brother of Raymond's mistress. Raymond and the brother fought, and Raymond was cut with a knife. Raymond pulled out a gun to shoot them, but Meursault took the gun from him and stated he would shoot them for Raymond. The two men ran away before Meursault was able to shoot. Later that afternoon, Meursault went walking on the beach by himself, still carrying Raymond's gun. He was not intentionally looking for the men again, but happened upon the man who had cut Raymond. The two stared at each other, hardly moving. Meursault knew he could just turn and leave, but he could not help but to move forward. The heat of the sun began bearing down on him, reminding Meursault of his mother's funeral procession. Before he was aware of what he did, Meursault pulled the trigger and shot the man. After a brief pause, Meursault fired four more times.

Macon

Macon is a married, middle-aged man who writes books about traveling to various parts of the world. He is the third son in a family of two brothers and one sister. Macon and his siblings were raised by their grandparents, after their widow mother, who had remarried, felt that she could no longer handle the children and a new husband. Macon's grandparents were supportive, caring individuals, who had a beautiful house and who insisted that the four grandchildren should learn to be quiet, thoughtful, and well-behaved. Order and intelligence were the major values in this family system. Overt expression of emotion was strongly de-valued. When viewed by others, this collection of four adults was characterized as eccentric individuals who did things like organizing their food pantry so that all of the cans were placed in alphabetical order. However, the three brothers and the sister saw themselves as a highly conventional, typical American family unit.

None of Macon's siblings were married. The two brothers lived with their sister, who inherited the grandparents' house and who organized the lives of her brothers. Macon went to college where he married a girl, Sarah who he had known in high school. Sarah and Macon had one child, Ethan and a dog named Edward. A major crisis in their life together occurred when their son was at summer camp. Ethan was shot and killed when entering a hamburger shop that was being robbed. Macon had no overt emotional response to the death of his son. However, images of his son kept intruding into his life. He had dreams of his son calling from camp and crying because no one had come to pick him up. Macon tried to protest to his son that they had been told he was dead.

Macon handled all of the funeral details following his son's death efficiently and was careful to be sure that the costs of the funeral were not unreasonable. He tried to be protective of his distraught wife. Almost a year later, Sarah left him because she felt that she was withering away in the context of Macon's emotional detachment from her. Macon tried to convince Sarah to stay, but did not get upset when she moved out.

Macon's closest friend after his wife moved out was his dog, Edward. While traveling, he boarded Edward at a facility named the Meow-Bow Animal Hospital. The dog, Edward became devoted to the woman who operated the facility. Through Edward, Macon and this woman, Muriel became friends. Eventually they became lovers, though Muriel was clearly the initiator in the relationship. As the relationship developed, Macon was aware of Muriel's interest in him, but his response was to be ambivalently disinterested or slightly amused. However, since his dog so clearly liked this flamboyant, outrageous woman who dressed in wildly colorful outfits, Macon kept seeing her.

Muriel had a son from a previous marriage. Macon found that interacting with the boy was both a painful and a positive healing process for him. Macon began to smile. He began to deviate from his carefully organized, staid existence. When Muriel suggested that they marry, Macon's ambivalence reappeared. He re-established contact with his former wife, Sarah who quickly and easily moved back into his life.

A second crisis in the family system occurred when Macon's sister, Rose met Macon's publisher. An immediate attraction developed between the two, and the publisher and the spinster sister decided to marry. This decision stimulated a crisis among Rose's brothers. At Thanksgiving dinner, Rose invited the publisher to eat with her brothers. The brothers attacked her cooking and refused to eat the turkey that she carefully labored over. The sister became verbally angry (a major taboo in this family system) and told her brothers that they were not going to doom her to becoming an old maid whose only reason for existence was to take care of them. The brothers sat silently while she ranted and stormed off. Then they each took a potato and ate that for their dinner. The publisher had two helpings of turkey. Six months later, the sister and the publisher were married.

Macon continued to travel and to write books with the goal of helping the traveler feel the least disruption from the effects of traveling. In his books, he talked about having a standard "carry-on" bag that contained all of the essentials for almost any place that one of his readers might need to visit. His descriptions of various cities focused on finding places to stay that were comfortable, lack intrusions, and permitted a traveler to move through efficiently. In effect, his goal was to help his readers go on trips without being impacted by the environments through which they might pass.

Muriel learned that Macon was going to Paris, and she booked a ticket on the same airline flight. Macon was upset by Muriel's presence and her persistence in trying to re-establish a relationship with him. This event also disrupted his work in Paris. And then, Macon hurt his back and ended up being bedridden. When Sarah learned of Macon's injury, she flew over to Paris to help him. She took over everything, including doing some of the work for his book. The effect of Sarah's controlling behavior on Macon helped him realize how he needed to develop more spontaneity in his life and how much he did love both Muriel and Muriel's son.

APPENDIX B: EXAMPLES OF ALEX'S WORD LISTS

Participant #62 (Alabama)

| | DSM | FFM | GAF |
|--------------------------|-----|-----|-----|
| borderline | 1 | 0 | 0 |
| sexually acting out | 1 | 1 | 0 |
| intelligent | 0 | 1 | 0 |
| attractive | 0 | 0 | 0 |
| homicidal | 0 | 0 | 0 |
| suicidal | 1 | 0 | 1 |
| sensitive to abandonment | 1 | 1 | 0 |
| emotionally volatile | 1 | 1 | 0 |
| sexually preoccupied | 1 | 1 | 0 |
| desperate | 0 | 0 | 0 |
| dramatic | 1 | 0 | 0 |
| no boundaries | 0 | 0 | 0 |
| manipulative | 0 | 1 | 0 |
| narcissistic | 1 | 0 | 0 |

Participant #81 (Idaho)

| Initial Task | DSM | FFM | GAF |
|-------------------------|-----|-----|-----|
| borderline personality | 1 | 0 | 0 |
| high functioning | 1 | 1 | 1 |
| push-pull relationships | 0 | 0 | 0 |
| need for attention | 1 | 0 | 0 |
| intelligent | 0 | 1 | 0 |
| Memory Cue | | | |
| Fatal Attraction | 0 | 0 | 0 |
| Follow Up Task | | | |
| high functioning | 1 | 1 | 1 |
| borderline | 1 | 0 | 0 |
| emotional | 1 | 1 | 0 |
| devaluation | 1 | 0 | 0 |
| idealizing | 1 | 0 | 0 |

Participant #115 (Connecticut)

| Initial Task | DSM | FFM | GAF |
|--|-----|-----|-----|
| flamboyant | 0 | 0 | 0 |
| histrionic | 1 | 0 | 0 |
| unstable | 1 | 1 | 0 |
| neurotic | 0 | 1 | 0 |
| borderline | 1 | 0 | 0 |
| cluster B | 1 | 0 | 0 |
| sexually unstable | 1 | 1 | 0 |
| uses sex to fill void | 1 | 0 | 0 |
| desperate | 0 | 0 | 0 |
| lacking self confidence | 1 | 1 | 0 |
| retaliatory | 0 | 0 | 0 |
| inflexible | 1 | 1 | 0 |
| dichotomous | 0 | 0 | 0 |
| impulsive | 1 | 1 | 0 |
| dependent | 1 | 1 | 0 |
| driven by fantasy | 1 | 1 | 0 |
| unable to self soothe | 0 | 0 | 0 |
| needy | 1 | 1 | 0 |
| Memory Cue | | | |
| borderline | 1 | 0 | 0 |
| Fatal Attraction | 0 | 0 | 0 |
| suicide for revenge | 1 | 0 | 1 |
| Follow Up Task | | | |
| borderline | 1 | 0 | 0 |
| vengeful | 0 | 0 | 0 |
| retaliatory | 0 | 0 | 0 |
| misuses sexuality | 1 | 0 | 0 |
| impulsive | 1 | 1 | 0 |
| history of sexual abuse | 0 | 1 | 0 |
| histrionic | 1 | 0 | 0 |
| bright | 0 | 0 | 0 |
| desperate | 0 | 0 | 0 |
| unable to modulate emotions | 1 | 1 | 0 |
| high/well-functioning in school/work tasks | 1 | 1 | 1 |
| mother had multiple marriages/divorces | 0 | 0 | 0 |
| poor attachment to male figures | 0 | 1 | 0 |
| fears abandonment | 1 | 1 | 0 |
| manipulative | 0 | 1 | 0 |
| emotionally unstable | 1 | 1 | 0 |
| reactive | 1 | 0 | 0 |
| uncontrolled in relationships | 1 | 1 | 0 |
| seeks romantic/sexual relationships to fill void | 1 | 0 | 0 |
| suicidal | 1 | 0 | 1 |

APPENDIX C: MATCHING TASK

The purpose of the matching task was to investigate the psychologists' ability to identify patients and their psychological presentations as described by the unique features of each classification system. Clinical utility was evaluated based on the psychologists' capacity to meaningfully and accurately interpret the clinical information offered from each respective system. Accuracy scores for the matching task were measured based on the frequency of the participants' correct answers for each case vignette across the three classification systems. The classification system receiving the highest level of accuracy for the matching task may suggest the properties of that system achieve greater clinical utility and value than the two alternative systems.

Pilot data were gathered during the current study to see how the participants ($n = 24$) would perform in this task. The small subset of participants received all nine case vignettes and a matching task worksheet for each classification system. Once the participants completed the word-list task and the *DSM*, FFM, and GAF score sheets for two case vignettes (as described in the Method I and II sections), the participants were instructed to complete a matching task with the remaining seven case vignettes. Participants read each of the seven case vignettes and were asked to match the case vignette with 12 *DSM* profiles, 12 FFM profiles, and 12 GAF profiles. The diagnostic profiles that the participants chose from were generated by the data from the Alabama sample in Method I, as well as three fabricated profiles, resulting in a total of 12 possible profile choices

for each classification system. For example, if Participant A received the case vignettes of Earnst and Alex for the word-list task and score sheet phase, then for the matching task Participant A matched Ted, Madeline, Charles, Aileen, Meursault, Marianne, and Macon to 12 profiles for each of the classification systems. The extra five profile choices (serving as decoys) consisted of the two case vignette profiles used in the word-list task and score sheet phase (in this example, Earnst and Alex) and three profiles that were created by the principle investigator to provide a greater breadth of choices. The choice of “None” was available to use as a response to the five profile choices not used.

The *DSM* profile choices consisted of the primary and secondary diagnoses for each case vignette, as determined by the Alabama sample. The FFM profile choices were based on the mean domain scores for each case vignette and were labeled according to the descriptions given to the values described on the FFM score sheet. For example, if a case vignette had a mean value of 4.79 for Neuroticism and 2.12 for Extraversion, these values were rounded to 5 and 2, respectively, and described as High Neuroticism and Problematic Low Extraversion. In addition to the labels, the mean domain scores were provided to better inform the participants about the severity of the domain described, thus maintaining the dimensional property of the FFM profiles. Domains with means rounded to 4.00 were not used as a description, because the case vignette was neither high nor low on that particular domain. For the GAF profile choices, each option consisted of the mean GAF score, as well as a range based on the mean lowest GAF score given and the mean highest GAF score given, as provided by the Alabama psychologists.

Accuracy scores were determined by tallying the correct matches. The scores for each classification model can range from 0 (no correct matches) to 12 (all correct

matches). The average accuracy scores for the *DSM* and FFM were 9.08 ($SD = 1.53$) and 5.75 ($SD = 2.80$), respectively. For the GAF, the accuracy score was 4.75 ($SD = 1.94$). A one-way, repeated measures ANOVA found significant mean differences in accuracy levels among the systems, $F(2, 46) = 30.41, p < .001$. Pairwise comparisons revealed that the accuracy score for the *DSM* was significantly higher than the FFM, $t(23) = 4.74, p < .001$, and the GAF, $t(23) = 10.01, p < .001$. No significant difference occurred in accuracy scores between the FFM and GAF. Based on this small subset of the sample, the *DSM* outperformed the FFM and GAF in terms of participants being able to match case vignettes to different classification profiles. This finding is not surprising given the participants' familiarity and extensive training with *DSM* diagnoses. More research is needed in the area of behavioral procedures to assess the clinical utility of the classification systems.

Table 1

Description of case vignettes

| Patient | Description | Word Count | Sentence Count |
|-----------|---|---------------|-------------------|
| Ted | Ted Bundy; serial killer who raped and murdered young, college women | 817 | 40 |
| Madeline | Prominent case study in Jerry Wiggins's (2003) <i>Paradigms of Personality Assessment</i> | 985 | 58 |
| Earnst | Historical case study in Henry A. Murray's (1938) <i>Explorations in Personality</i> | 890 | 42 |
| Charles | Charles Guiteau; assassin of President James Garfield | 1014 | 56 |
| Aileen | Aileen Wuornos; serial killer and female character in the movie <i>Monster</i> | 1024 | 65 |
| Meursault | Main male character in the novel <i>The Stranger</i> by Albert Camus | 1036 | 63 |
| Marianne | Main female character in the movie <i>Zuckerbaby (Sugarbaby)</i> | 993 | 57 |
| Alex | Alex Forrest; main female character in the movie <i>Fatal Attraction</i> | 1048 | 63 |
| Macon | Macon Leary; main male character in the novel <i>The Accidental Tourist</i> by Anne Tyler | 901 | 56 |

Table 2

Percentage of *DSM* diagnoses for each case vignette

| | Ted (<i>n</i> = 36) | Madeline (<i>n</i> = 31) | Earnst (<i>n</i> = 32) | Charles (<i>n</i> = 35) | Aileen (<i>n</i> = 41) | Meursault (<i>n</i> = 37) | Marianne (<i>n</i> = 31) | Alex (<i>n</i> = 36) | Macon (<i>n</i> = 41) |
|--------------|-------------------------|------------------------------|----------------------------|-----------------------------|----------------------------|-------------------------------|------------------------------|--------------------------|---------------------------|
| Paranoid | 3.2% | | 7.1% | 13.3% | 10.8% | | | 6.9% | |
| Schizoid | 22.6% | | 53.6% | 6.7% | 5.4% | 53.3% | 21.4% | | 32.1% |
| Schizotypal | 3.2% | | 14.3% | 58.6% | 5.4% | | 17.9% | | |
| Antisocial | 83.9% | 20.7% | | 23.3% | 81.1% | 13.3% | | 3.4% | |
| Borderline | 3.2% | 10.3% | | 13.3% | 48.6% | | 25.0% | 93.1% | 3.6% |
| Histrionic | | 62.1% | | 3.3% | 21.6% | | 3.6% | 24.1% | |
| Narcissistic | 48.4% | 51.7% | 3.6% | 56.7% | 24.3% | 10.0% | 3.6% | 17.2% | 3.6% |
| Avoidant | 3.2% | | 67.9% | 3.3% | | 3.3% | 14.3% | | 7.1% |
| Dependent | | | | | | | 3.6% | 17.2% | 3.6% |
| OCPD | | | 10.7% | 3.3% | 2.8% | 3.3% | 7.1% | | 35.7% |
| NOS | 16.1% | 24.1% | 17.9% | 16.7% | 24.3% | 30.0% | 50.0% | 6.9% | 21.4% |

Note. Percentages not shown indicate that no personality disorder was diagnosed for the case vignette.

Table 3

Mean (*SD*) *DSM* prototypicality ratings for each case vignette

| | Ted (<i>n</i> = 36) | Madeline (<i>n</i> = 31) | Earnst (<i>n</i> = 32) | Charles (<i>n</i> = 35) | Aileen (<i>n</i> = 41) | Meursault (<i>n</i> = 37) | Marianne (<i>n</i> = 31) | Alex (<i>n</i> = 36) | Macon (<i>n</i> = 41) |
|--------------|-------------------------|------------------------------|----------------------------|-----------------------------|----------------------------|-------------------------------|------------------------------|--------------------------|---------------------------|
| Paranoid | 2.17 (1.45) | | 2.74 (1.29) | 3.06 (2.02) | 2.79 (1.56) | | | 2.66 (1.52) | |
| Schizoid | 3.21 (1.74) | | 5.00 (1.69) | 2.88 (1.49) | 1.95 (1.18) | 4.84 (1.84) | 3.70 (1.60) | | 3.95 (1.78) |
| Schizotypal | 1.86 (1.48) | | 3.00 (1.59) | 5.01 (1.92) | 2.11 (1.41) | | 2.81 (1.84) | | |
| Antisocial | 6.29 (0.76) | 3.53 (1.38) | | 3.74 (1.51) | 6.07 (1.31) | 3.73 (1.72) | | 3.14 (1.32) | |
| Borderline | 1.88 (1.12) | 2.98 (1.67) | | 2.73 (1.60) | 4.55 (1.78) | | 3.55 (1.62) | 6.56 (0.56) | 1.41 (0.85) |
| Histrionic | | 5.58 (1.12) | | 2.43 (1.25) | 3.54 (1.86) | | 2.30 (1.29) | 4.64 (1.38) | |
| Narcissistic | 4.85 (1.81) | 5.47 (1.31) | 2.16 (1.37) | 4.91 (1.74) | 3.64 (1.74) | 3.01 (1.86) | 1.54 (1.08) | 3.68 (1.44) | 1.64 (1.14) |
| Avoidant | 2.02 (1.36) | | 4.74 (1.73) | 1.79 (0.98) | | 1.99 (1.67) | 3.48 (1.68) | | 2.99 (1.68) |
| Dependent | | | | | | | 2.10 (1.42) | 3.34 (1.78) | 2.57 (1.51) |
| OCPD | | | 2.37 (1.56) | 1.86 (1.07) | 1.39 (0.97) | 1.78 (1.26) | 2.62 (1.72) | | 4.34 (1.74) |
| NOS | 2.70 (2.15) | 2.93 (2.25) | 3.10 (1.85) | 3.24 (1.99) | 2.83 (2.18) | 2.97 (2.11) | 4.52 (1.97) | 2.83 (1.95) | 3.00 (2.16) |

Note. Ratings not shown indicate that no personality disorder was diagnosed for the case vignette.

Table 4

Mean (*SD*) FFM ratings for each case vignette

| | Ted (<i>n</i> = 36) | Madeline (<i>n</i> = 31) | Earnst (<i>n</i> = 32) | Charles (<i>n</i> = 35) | Aileen (<i>n</i> = 41) | Meursault (<i>n</i> = 37) | Marianne (<i>n</i> = 31) | Alex (<i>n</i> = 36) | Macon (<i>n</i> = 41) |
|---------------------|-------------------------|------------------------------|----------------------------|-----------------------------|----------------------------|-------------------------------|------------------------------|--------------------------|---------------------------|
| <i>Neuroticism</i> | 2.67 (0.78) | 3.01 (0.78) | 4.76 (0.57) | 4.30 (1.10) | 4.68 (0.81) | 3.26 (0.81) | 4.84 (0.78) | 5.25 (0.64) | 3.40 (0.73) |
| Anxiousness | 1.64 (0.93) | 2.67 (1.37) | 5.94 (1.05) | | | 1.65 (0.98) | 4.74 (1.29) | 5.06 (1.27) | |
| Angry | 5.00 (1.90) | | | 5.97 (1.13) | 6.68 (0.57) | | 4.61 (1.54) | 6.53 (0.65) | 2.58 (1.15) |
| Depressiveness | 2.41 (1.23) | 2.63 (0.96) | 5.35 (0.98) | | 5.03 (1.35) | | 5.06 (1.03) | 5.00 (1.15) | |
| Self-conscious | 1.47 (0.70) | 1.52 (0.78) | 5.84 (1.02) | 2.52 (1.64) | 2.40 (1.39) | 2.50 (1.18) | 5.13 (1.15) | 2.97 (1.21) | |
| Impulsivity | | 4.83 (1.80) | 2.37 (0.85) | 5.86 (1.50) | 6.38 (1.21) | 4.58 (1.65) | | 6.31 (1.14) | 1.90 (1.01) |
| Vulnerability | 1.69 (0.80) | 2.23 (0.92) | | 3.34 (1.62) | 3.38 (1.46) | 2.91 (1.25) | 5.00 (0.82) | 5.61 (1.02) | |
| <i>Extraversion</i> | 4.77 (0.86) | 5.89 (0.57) | 2.12 (0.58) | 4.41 (0.80) | 4.71 (1.00) | 2.49 (0.94) | 2.59 (0.94) | 5.45 (0.67) | 2.46 (0.66) |
| Warmth | 1.73 (1.07) | | 2.23 (1.14) | 2.03 (0.95) | 2.10 (1.00) | 1.14 (0.42) | 2.77 (1.52) | 4.74 (1.20) | 2.05 (0.89) |
| Gregariousness | 4.72 (1.78) | 6.48 (0.77) | 1.72 (0.63) | 3.18 (1.73) | | 2.16 (1.28) | 1.87 (1.02) | 5.56 (0.88) | 2.49 (1.08) |
| Assertiveness | 5.86 (1.17) | 6.33 (0.66) | 2.06 (1.03) | 5.66 (1.28) | 5.59 (1.30) | 3.35 (1.57) | 3.10 (1.30) | 5.64 (0.87) | 2.54 (0.90) |
| Activity | 5.71 (1.07) | 6.42 (0.67) | 2.68 (1.38) | 5.71 (1.27) | 5.56 (1.38) | 2.73 (1.24) | 2.58 (1.20) | 5.31 (0.93) | 3.25 (1.06) |
| Excitement | 5.83 (1.60) | 6.40 (0.88) | 2.00 (0.82) | 5.55 (1.15) | 6.34 (1.35) | 3.22 (1.65) | 2.84 (1.37) | 6.17 (0.94) | 2.20 (0.78) |
| Positive | | 5.94 (0.85) | 2.10 (0.75) | | 4.54 (1.68) | 2.35 (1.06) | 2.35 (1.05) | 5.25 (1.18) | 2.25 (0.78) |

Table 4 (continued)

| | Ted (<i>n</i> = 36) | Madeline (<i>n</i> = 31) | Earnst (<i>n</i> = 32) | Charles (<i>n</i> = 35) | Aileen (<i>n</i> = 41) | Meursault (<i>n</i> = 37) | Marianne (<i>n</i> = 31) | Alex (<i>n</i> = 36) | Macon (<i>n</i> = 41) |
|----------------------|-------------------------|------------------------------|----------------------------|-----------------------------|----------------------------|-------------------------------|------------------------------|--------------------------|---------------------------|
| <i>Openness</i> | 4.32 (0.94) | 4.68 (0.76) | 3.33 (0.86) | 5.07 (0.87) | 4.76 (1.17) | 2.78 (0.86) | 4.28 (1.22) | 5.40 (0.72) | 2.46 (0.65) |
| Fantasy | | | 4.81 (1.75) | 6.46 (0.98) | 4.85 (1.50) | 2.76 (1.46) | 5.65 (1.70) | 5.69 (0.99) | 2.29 (1.15) |
| Aesthetic | | | | 4.78 (1.74) | 3.29 (1.74) | 2.00 (1.15) | | 4.74 (0.93) | 2.98 (1.17) |
| Feelings | 2.31 (1.20) | | 2.97 (1.65) | | | 1.35 (0.54) | | 5.60 (1.59) | 2.02 (0.99) |
| Actions | 5.11 (1.39) | 5.48 (1.18) | 2.35 (1.28) | 6.29 (1.20) | 5.95 (1.47) | | | 6.22 (0.93) | 2.24 (0.94) |
| Ideas | 5.38 (1.48) | 4.62 (1.05) | 3.37 (1.54) | 6.40 (1.40) | 5.38 (1.37) | 3.43 (1.82) | 5.45 (1.65) | 5.20 (1.05) | 2.49 (1.10) |
| Values | | 5.63 (0.89) | 2.62 (1.21) | 2.39 (1.58) | 5.03 (1.71) | 3.49 (1.40) | | 4.76 (1.25) | 2.81 (0.91) |
| <i>Agreeableness</i> | 1.70 (0.48) | 2.49 (0.64) | 4.08 (0.80) | 2.33 (0.79) | 1.63 (0.52) | 3.41 (1.00) | 4.09 (0.68) | 2.79 (0.90) | 4.36 (0.63) |
| Trust | 2.30 (1.05) | 3.48 (1.18) | 2.61 (1.54) | 2.39 (1.22) | 1.80 (0.71) | 3.47 (1.31) | | | |
| Straightforward | 1.14 (0.42) | 2.52 (1.34) | 4.69 (0.93) | 3.03 (1.61) | 1.61 (0.86) | | | 2.34 (1.53) | 4.62 (0.99) |
| Altruism | 1.15 (0.36) | 2.77 (1.18) | | 2.61 (1.38) | 1.68 (0.82) | 2.72 (1.37) | | 2.39 (1.27) | |
| Compliance | 2.34 (1.41) | 2.38 (0.86) | 4.50 (1.11) | 2.00 (1.00) | 1.34 (0.57) | | | 2.23 (1.24) | 4.90 (0.83) |
| Modesty | 1.92 (1.18) | 1.65 (0.61) | 4.97 (1.22) | 1.71 (1.03) | 2.02 (1.01) | | 4.72 (1.10) | 3.21 (0.95) | 4.68 (1.02) |
| Tender-minded | 1.33 (0.53) | 2.03 (0.91) | | 2.27 (1.07) | 1.34 (0.57) | 2.09 (0.98) | | 3.09 (1.29) | |

Table 4 (continued)

| | Ted (n = 36) | Madeline (n = 31) | Ernst (n = 32) | Charles (n = 35) | Aileen (n = 41) | Meursault (n = 37) | Marianne (n = 31) | Alex (n = 36) | Macon (n = 41) |
|----------------------|--------------------|----------------------|-------------------|---------------------|--------------------|-----------------------|----------------------|------------------|--------------------|
| <i>Conscientious</i> | 4.87 (1.14) | 4.77 (0.75) | 4.80 (0.95) | 2.71 (1.24) | 2.41 (0.98) | 3.52 (0.95) | 4.56 (0.93) | 3.69 (0.73) | 5.51 (0.65) |
| Competence | 5.65 (1.20) | 5.68 (0.94) | 4.83 (1.39) | 2.49 (1.46) | 3.23 (1.56) | | | 4.54 (0.66) | 5.80 (0.94) |
| Order | 5.82 (1.00) | 4.78 (1.01) | 4.86 (0.99) | 2.24 (1.46) | 2.92 (1.35) | 4.56 (1.05) | 5.13 (1.22) | | 6.07 (0.88) |
| Dutifulness | | | 4.86 (0.97) | 2.16 (1.35) | 1.95 (1.22) | | 4.77 (1.12) | 2.97 (1.03) | 5.63 (0.73) |
| Achievement | 5.63 (0.87) | 6.29 (0.74) | 4.72 (1.49) | | 2.48 (1.18) | | | 4.74 (0.96) | 5.13 (0.91) |
| Self-discipline | | | 4.81 (1.14) | 3.11 (1.60) | 2.22 (1.27) | 3.19 (1.24) | 4.52 (1.12) | | 5.26 (0.72) |
| Deliberation | 4.60 (1.67) | | 5.00 (1.13) | 2.56 (1.80) | 1.80 (1.14) | 2.57 (1.30) | 4.65 (1.62) | 2.69 (1.53) | 5.18 (0.93) |

Note. Ratings rounded to 4.00 were deleted, except for domain ratings.

Table 5

GAF profiles for each case vignette

| | Ted (<i>n</i> = 36) | Madeline (<i>n</i> = 31) | Earnst (<i>n</i> = 32) | Charles (<i>n</i> = 35) | Aileen (<i>n</i> = 41) | Meursault (<i>n</i> = 37) | Marianne (<i>n</i> = 31) | Alex (<i>n</i> = 36) | Macon (<i>n</i> = 41) |
|-----------|-------------------------|------------------------------|----------------------------|-----------------------------|----------------------------|-------------------------------|------------------------------|--------------------------|---------------------------|
| GAF score | 17.97 | 65.16 | 51.75 | 19.16 | 20.32 | 42.39 | 36.65 | 27.03 | 62.73 |
| <i>SD</i> | 18.42 | 14.40 | 7.87 | 12.47 | 16.71 | 19.94 | 12.58 | 18.44 | 9.02 |
| range | 1-75 | 35-95 | 33-65 | 1-51 | 1-60 | 1-70 | 10-60 | 1-62 | 45-85 |
| Low GAF | 10.49 | 57.97 | 43.98 | 10.97 | 12.98 | 32.15 | 24.18 | 16.08 | 54.62 |
| <i>SD</i> | 15.08 | 13.05 | 9.63 | 11.32 | 14.12 | 20.46 | 13.73 | 15.75 | 11.09 |
| range | 1-55 | 35-91 | 30-60 | 1-51 | 1-51 | 0-65 | 1-52 | 1-55 | 31-81 |
| High GAF | 30.44 | 71.90 | 58.88 | 34.06 | 29.54 | 52.53 | 44.66 | 39.75 | 69.28 |
| <i>SD</i> | 26.39 | 14.41 | 8.58 | 15.19 | 17.26 | 19.13 | 13.09 | 19.55 | 8.53 |
| range | 1-89 | 40-100 | 37-80 | 5-60 | 1-60 | 10-80 | 15-65 | 9-70 | 50-90 |

Table 6

Reliability coefficients for the *DSM* prototypicality ratings and FFM facet ratings

| | <i>n</i> | Cronbach's Alpha | |
|-----------|----------|------------------|------|
| | | <i>DSM</i> | FFM |
| Ted | 36 | .982 | .984 |
| Madeline | 31 | .987 | .984 |
| Earnst | 32 | .966 | .976 |
| Charles | 35 | .962 | .977 |
| Aileen | 41 | .983 | .987 |
| Meursault | 37 | .964 | .947 |
| Marianne | 31 | .944 | .939 |
| Alex | 36 | .987 | .981 |
| Macon | 41 | .974 | .987 |

Note. Missing values replaced with means.

Table 7

Mean (*SD*) utility ratings of the *DSM*, *FFM*, and *GAF* for each case vignette

| | Confident | | | Usefulness | | |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| | <i>DSM</i> | <i>FFM</i> | <i>GAF</i> | <i>DSM</i> | <i>FFM</i> | <i>GAF</i> |
| Total | 5.22 ^a | 4.42 ^b | 4.83 ^c | 4.84 ^a | 4.58 ^a | 3.80 ^b |
| (<i>n</i> = 320) | (1.06) | (1.01) | (1.27) | (1.38) | (1.19) | (1.54) |
| Ted | 5.53 | 4.33 | 4.80 | 5.26 | 4.67 | 3.41 |
| (<i>n</i> = 36) | (1.13) | (1.15) | (1.62) | (1.34) | (1.33) | (1.73) |
| Madeline | 5.23 | 4.48 | 4.65 | 5.06 | 4.85 | 3.60 |
| (<i>n</i> = 31) | (0.99) | (0.85) | (1.40) | (1.46) | (1.10) | (1.63) |
| Earnst | 4.94 | 4.34 | 4.84 | 4.84 | 4.50 | 3.84 |
| (<i>n</i> = 32) | (1.03) | (1.15) | (1.08) | (1.19) | (1.37) | (1.37) |
| Charles | 4.99 | 4.43 | 5.23 | 4.74 | 4.26 | 4.57 |
| (<i>n</i> = 35) | (1.00) | (0.98) | (1.11) | (1.40) | (1.22) | (1.40) |
| Aileen | 5.38 | 4.88 | 4.88 | 4.95 | 5.02 | 3.68 |
| (<i>n</i> = 41) | (1.08) | (1.00) | (1.38) | (1.32) | (1.11) | (1.47) |
| Meursault | 4.97 | 4.13 | 4.32 | 4.39 | 4.28 | 3.27 |
| (<i>n</i> = 37) | (1.00) | (0.88) | (1.30) | (1.39) | (1.11) | (1.47) |
| Marianne | 4.64 | 3.97 | 4.73 | 4.12 | 4.26 | 4.23 |
| (<i>n</i> = 31) | (1.04) | (1.08) | (0.91) | (1.46) | (1.06) | (1.36) |
| Alex | 6.00 | 4.66 | 4.87 | 5.66 | 4.49 | 3.73 |
| (<i>n</i> = 36) | (0.84) | (0.80) | (1.38) | (1.03) | (1.01) | (1.50) |
| Macon | 5.19 | 4.46 | 5.10 | 4.49 | 4.74 | 3.94 |
| (<i>n</i> = 41) | (0.90) | (0.93) | (0.98) | (1.27) | (1.22) | (1.63) |

Note. Means under the same headings of “Confident” and “Usefulness” that do not share superscripts differ at $p < .001$.

Table 8

Mean (*SD*) absolute differences for each case vignette across the *DSM* (based on the mean of two most frequent diagnoses), FFM (based on the mean of the five domains), and GAF (based on the range) on a 7-point rating scale

| | <i>DSM</i> | FFM | GAF | Total |
|----------------------------|--------------------------|--------------------------|--------------------------|-------------|
| Ted (<i>n</i> = 36) | 1.09 (0.53) ^a | 0.65 (0.31) ^b | 0.83 (0.67) | 0.86 (0.33) |
| Madeline (<i>n</i> = 31) | 0.97 (0.44) ^a | 0.55 (0.23) ^b | 0.27 (0.20) ^c | 0.60 (0.19) |
| Earnst (<i>n</i> = 32) | 1.44 (0.62) ^a | 0.57 (0.26) ^b | 0.33 (0.26) ^c | 0.78 (0.24) |
| Charles (<i>n</i> = 35) | 1.49 (0.83) ^a | 0.74 (0.34) ^b | 0.72 (0.48) ^b | 0.98 (0.38) |
| Aileen (<i>n</i> = 41) | 1.18 (0.71) ^a | 0.68 (0.35) ^b | 0.54 (0.42) ^b | 0.80 (0.28) |
| Meursault (<i>n</i> = 37) | 1.74 (0.76) ^a | 0.72 (0.34) ^b | 0.67 (0.52) ^b | 1.03 (0.32) |
| Marianne (<i>n</i> = 31) | 1.52 (0.66) ^a | 0.72 (0.27) ^b | 0.60 (0.52) ^b | 0.94 (0.31) |
| Alex (<i>n</i> = 36) | 0.83 (0.38) ^a | 0.56 (0.27) ^b | 0.70 (0.51) | 0.70 (0.21) |
| Macon (<i>n</i> = 41) | 1.47 (0.64) ^a | 0.52 (0.22) ^b | 0.38 (0.32) ^c | 0.78 (0.28) |
| Total (<i>n</i> = 320) | 1.30 (0.69) ^a | 0.63 (0.30) ^b | 0.56 (0.49) ^c | 0.83 (0.31) |

Note. Means for each case vignette that do not share superscripts differ at $p < .001$,

except for:

Alex— $p < .01$

Macon—FFM/GAF, $p < .05$

Total—FFM/GAF, $p < .05$

Table 9

Ranking of variability values for each case vignette across the classification systems

| Rank | DSM | FFM | GAF | Total |
|------|-----------|------------------------|-----------|------------------|
| 1 | Alex | Macon | Madeline | Madeline |
| 2 | Madeline | Madeline | Earnst | Alex |
| 3 | Ted | Alex | Macon | Earnst and Macon |
| 4 | Aileen | Earnst | Aileen | |
| 5 | Earnst | Ted | Marianne | Aileen |
| 6 | Macon | Aileen | Meursault | Ted |
| 7 | Charles | Meursault and Marianne | Alex | Marianne |
| 8 | Marianne | | Charles | Charles |
| 9 | Meursault | Charles | Ted | Meursault |

Note. A ranking of 1 represents best agreement (low variability value). A ranking of 9 represents worst agreement (high variability value).

Table 10

Frequency (and percentage) of descriptors across all case vignettes for each classification system

| | Alabama Sample (<i>n</i> = 71) | Idaho/Connecticut Sample (<i>n</i> = 44) | | | Combined Sample (<i>n</i> = 115) | |
|-------------------|---|--|--------------|------------------|---|---------------|
| | | First Word-List | Memory Cues | Second Word-List | Total | |
| <i>DSM</i> | 691 (45.73%) | 480 (49.43%) | 160 (44.08%) | 289 (50.70%) | 929 (48.79%) | 1620 (47.44%) |
| FFM | 781 (51.69%) | 528 (54.38%) | 104 (28.65%) | 273 (47.89%) | 905 (47.53%) | 1686 (49.37%) |
| GAF | 114 (7.54%) | 76 (7.83%) | 36 (9.92%) | 41 (7.19%) | 153 (8.04%) | 267 (7.82%) |
| Total | 1511 | 971 | 363 | 570 | 1904 | 3415 |
| <i>Pure DSM</i> | 234 (40.84%) | 148 (40.33%) | 76 (66.09%) | 113 (50.90%) | 337 (47.87%) | 571 (44.71%) |
| <i>Pure FFM</i> | 338 (58.99%) | 214 (58.31%) | 35 (30.43%) | 108 (48.65%) | 357 (50.71%) | 695 (54.52%) |
| <i>Pure GAF</i> | 1 (0.17%) | 5 (1.36%) | 4 (3.38%) | 1 (0.45%) | 10 (1.42%) | 11 (0.86%) |
| <i>Pure Total</i> | 573 | 367 | 115 | 222 | 704 | 1277 |