2006

Diagnosing Borderline Personality Disorder: The Effect Of Therapists' Negative Emotional Reactions On Diagnostic Judgments

Keith Mayo
Indiana State University

Keith Wyatt Mayo

Education

Currently:	Psy.D. Student, Doctoral Program in Clinical Psychology *Indiana State University* Anticipated Graduation Date: 12/06
2003	Master of Science in Psychology *Indiana State University*
1998	Bachelor of Science in Psychology with Honors, May 1998 *Middle Tennessee State University*

Supervised Clinical Experience

September 2005 – September 2006	Psychology Intern (*full time, salaried*) *Pre-Doctoral Internship – APA Accredited* *Veterans Affairs Medical Center, Lexington, KY*
September 2004 – May 2005	Student Therapist (hourly) *Hamilton Center, Inc., Clinton, IN*
July 2003 – May 2004	Student Therapist (hourly) *Hamilton Center Inc., Terre Haute, IN*
August 2002 – August 2004	Graduate Student Therapist (unpaid) *Indiana State Psychology Clinic, Terre Haute, IN*

Research Experience

Dissertation: "Diagnosing borderline personality disorder: The effect of negative emotional response on diagnostic judgments." (Proposal approved October, 2004).

August 2001 – May 2002
Research Assistant
Department of Psychology, Indiana State University
Assisted Dr. June Sprock with data collection, analysis, and web page construction. Met weekly with Dr. Sprock's research team to discuss current research efforts and student projects.

Teaching Experience

August 2002 – May 2003
Graduate Student Instructor
Indiana State University Psychology Department

DIAGNOSING BORDERLINE PERSONALITY DISORDER: THE EFFECT OF THERAPISTS' NEGATIVE EMOTIONAL REACTIONS ON DIAGNOSTIC JUDGMENTS

A Dissertation

Presented to

The School of Graduate Studies

Department of Psychology

Indiana State University

Terre Haute, Indiana

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Psychology

---

by

Keith Mayo

December 2006

School of Graduate Studies
Indiana State University
Terre Haute, Indiana

CERTIFICATE OF APPROVAL

DOCTORAL DISSERTATION

This is to certify that the Doctoral Dissertation of

Keith W. Mayo

entitled

Diagnosing Borderline Personality Disorder: The Effect of Therapists' Negative Emotional Reactions on Diagnostic Judgments

has been approved by the Examining Committee for the dissertation requirement for the

Doctor of Psychology degree

in Clinical Psychology

December 2006

June Sprock, Ph.D. Date
CHAIR, Dissertation Committee

Michael J. Murphy, Ph.D., ABPP Date
Member, Dissertation Committee

Veanne Anderson, Ph.D. Date
Member, Dissertation Committee

Jolynn S. Kuhlman, Ph.D. Date
Interim Dean, School of Graduate Studies

ABSTRACT

Previous studies suggest that clinicians are prone to bias in diagnosing Borderline Personality Disorder (BPD) and that BPD symptoms elicit negative emotional reactions (NER) from clinicians. However, no studies have specifically examined the effect of NER on the diagnosis of BPD. This study examined the decision-making processes used when assigning a diagnosis of BPD, specifically, whether clinicians' NER towards patients exhibiting BPD symptoms bias decision-making and result in misuse of the BPD diagnosis. A randomly-selected national sample of 98 licensed psychologists completed an Internet survey in which they read two case vignettes that were designed to elicit NER but were below threshold for a diagnosis of BPD. Participants rated the representativeness of a series of Axis I and II diagnoses for the case; assigned Axis I and II diagnoses and rated their level of confidence; rated severity, prognosis, and the likelihood of the individual in the case benefiting from treatment; and rated the applicability of a series of symptoms for the case (including each of the DSM-IV criteria for BPD). They then rated the degree of NER felt toward the patient using two subscales of the Impact Message Inventory (IMI). Results provided moderate support for the prediction that participants who report higher levels of NER would be more likely to diagnose BPD, would assign higher BPD representativeness ratings, and would rate the prognosis and likelihood of response to treatment lower. Predictions concerning the moderating effects of clinician variables (years of clinical experience, percentage of time

spent in direct patient contact) were not supported, but clinician gender had significant effects on the diagnosis of BPD. The hypothesis that clinicians who were asked to assign diagnoses before rating symptoms (i.e., a simulated prototype approach) would be more prone to over-diagnosis of BPD was also not supported, but order of the cases had unexpected effects on the results. Implications for clinical training and directions for future research are discussed.

ACKNOWLEDGMENTS

I would like to thank Dr. June Sprock, whose dedication to this project and patient guidance I will always remember. I would also like to acknowledge the support of my partner, Jamey Smith, and my parents Jerry and Lynda Mayo.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 1

INTRODUCTION

Overview

The way in which clinicians use the diagnosis of Borderline Personality Disorder (BPD) has been criticized in a number of ways. Previous studies (e.g., Becker & Lamb, 1994; Blashfield & Herkov, 1996; Morey and Ochoa, 1989; Ochoa & Morey, 1990) have shown that clinicians do not weigh BPD symptoms equally in formulating the diagnosis and may be more prone to diagnose BPD in women than in men. In addition, these studies show that clinician characteristics may also affect the diagnosis of BPD. Female clinicians may over-diagnose BPD relative to their male counterparts, psychiatrists and psychologists weigh BPD criteria differently, clinicians with fewer years of experience diagnose BPD more readily, and clinicians whose work involves more direct patient contact hours diagnose BPD more frequently. As with many of the personality disorders, the criteria for BPD are notoriously vague, and the above findings may be a reflection of the ambiguity and heterogeneity of the criteria, a problem that has been discussed extensively in previous reviews (e.g., Gunderson, Zanarini, & Kisiel, 1995). Other studies point to the intense negative emotional reactions that clinicians and other treatment professionals exhibit towards patients diagnosed with BPD (e.g., McIntyre & Schwartz, 1998). These studies suggest that another, previously unstudied factor may

lead to a misuse of the BPD diagnosis: the negative emotional reactions that clinicians exhibit toward problematic patients.

The literature examining the process of human decision making provides support for the notion that negative emotional reactions (NER) may contribute to inaccurate diagnosis. Studies (e.g., Cumming & Harris, 2001; Edland & Svenson, 1993; Orasanu, 1997) have illustrated that under conditions of stress, anxiety, or heightened emotional arousal, individuals are more prone to use decision-making "shortcuts" (i.e., heuristics) whereby they attend only to aspects of the task that are viewed as important and ignore other, potentially relevant, information. When this occurs in the context of medical decision making, clinicians may ignore information such as the base rates of a diagnosis in a population and fail to question the reliability of the information that is presented to them (Hall, 2002). In addition, both positive and negative affect have been shown to have complex effects on decision-making biases (e.g., Jundt & Hinsz, 2002). Thus, NER towards patients may lead to unscientific decision-making habits that are based on a superficial understanding of relevant data, and influence the process of assigning an accurate diagnosis.

The present study investigated the influence of NER on diagnostic decision making with respect to BPD. The behaviors embodied in the diagnostic criteria for BPD (e.g., impulsive anger, suicidal gestures, intense interpersonal relationships) have a high potential for producing NER in others, which may negatively influence diagnostic decision making. As a result, the accuracy of their diagnoses would suffer, and patients with subthreshold symptoms might inappropriately receive a diagnosis of BPD. Since BPD is quickly becoming a diagnosis of denigration characterized by overuse and misuse

(Becker & Lamb, 1994), it becomes important to investigate the processes that underlie this trend. However, no published studies have specifically examined the role of NER in diagnostic decision-making, or more directly, the role of NER in formulating a diagnosis of BPD. Therefore, the present study was designed to examine the influence of NER, a previously overlooked but potentially important variable, on the diagnosis and over-diagnosis of BPD. More generally, this study will help increase our understanding of clinician bias and diagnostic decision-making.

Overview of Borderline Personality Disorder (BPD) Diagnosis

According to the American Psychiatric Association's (1994) *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition, BPD is primarily marked by recurrent interpersonal difficulties, affective instability, and difficulties maintaining a consistent self-image. Impulsive behaviors, such as inappropriate displays of anger and suicidal gestures (e.g., self-mutilation) are also core features of BPD. The specific criteria for BPD, outlined in Appendix A, contain a heterogeneous mixture of affective, cognitive, and behavioral symptoms that frequently occur in response to interpersonal distress. In essence, patients diagnosed with BPD are very sensitive to changes in interpersonal relationships, especially those that trigger feelings of abandonment, and react with intense cognitive and emotional dysregulation when such changes occur. Other features of this diagnosis include chronic feelings of emptiness and brief psychosis and/or dissociation in response to stress. According to the DSM-IV, BPD is estimated to have a prevalence of approximately 2% of the general population, 10% of those seen in outpatient clinics, and 20% of those being treated on inpatient units. Thus, patients diagnosed with BPD place a significant toll on the mental healthcare system

(Lenzenweger & Cicchetti, 2005). In fact, recent research (Sansone, Songer, & Miller, 2005) has shown that individuals with BPD have significantly higher rates of healthcare utilization (e.g., longer inpatient stays, more frequent outpatient treatment) as well as more frequent episodes of self-harm behavior when compared to patients without the disorder.

Problems with the BPD Diagnosis

Diagnostic Criteria. As with many of the Axis II disorders, the criteria for BPD have been widely criticized. Since the introduction of BPD in the DSM-III (APA, 1980) the criteria for BPD have been criticized as heterogeneous, non-specific, and thus more likely to be applied in an inaccurate, or even biased, manner. In a meta-analytic review conducted by Gunderson, Zanarini, and Kisiel (1991), inherently vague DSM-III-R (1987) diagnostic criteria such as "identity disturbance" and "affective instability" were shown to lead to high rates of comorbidity with both Axis I and Axis II disorders. These criteria received only minor modifications in the DSM-IV (APA, 1994), and Taylor (1995) pointed out that these modifications may have led to these criteria becoming even more imprecise and overlapping more with other disorders. Moreover, recent research has shown that specific behaviors, such as acts of self-harm and impulsive behavior, may be given more weight by clinicians in determining a diagnosis of BPD (Burgmer, Jessen, & Freyberger, 2000). This *form of bias appears to be especially problematic when clinicians do not use structured interviews to diagnose BPD (Sansone, Songer, & Gaither, 2001).* Further, studies by Blashfield and Herkov (1996) and Morey and Ochoa (1989) show that the presence of these symptoms lead clinicians to diagnose BPD in

cases in which there are not enough symptoms to warrant the diagnosis (i.e., over-diagnosis).

Gender bias. One area that has received considerable attention and which is relevant to this study is gender bias in the diagnosis of BPD. The research addressing this issue (e.g., Becker & Lamb, 1994) has demonstrated that clinicians are influenced by factors other than the patient's presenting symptoms when assigning a diagnosis of BPD, and illustrate the potential for a BPD diagnosis to be misused by clinicians in response to attitudinal biases. At the heart of this issue is the disproportional number of women who receive the diagnosis relative to men. Widiger and Trull (1993) reviewed 75 non-overlapping studies on BPD prevalence rates and found that, on average, women accounted for 76% of the patients who received a diagnosis of BPD. This finding, which is nearly identical to the 75% prevalence rate for women reported in the DSM-IV (APA, 1994), begs the question of sex bias for this diagnosis and has raised the suspicion of a number of researchers. Previous researchers have explained the disproportional number of women receiving the BPD diagnosis by arguing that clinicians may inappropriately use gender stereotypes to guide their assessment of pathological behavior. Men who exhibit excessive displays of anger may not be viewed as pathological as this behavior is consistent with their gender stereotype (Nehls, 1998). Moreover, Zoltnick, Rothschild, and Zimmerman (2002) argued that men may display emotional instability in a more outwardly directed manner than women, thus leading to a diagnosis of Antisocial Personality Disorder as opposed to BPD.

However, differential prevalence of a disorder is not sufficient evidence of diagnostic bias (Widiger & Spitzer, 1991; Widiger, 1998). Therefore, researchers have

used a number of different methodologies to examine diagnostic bias. One of the most widely used methods is to present clinical vignettes in which the variable of interest (i.e., patient sex) is manipulated, while all other variables are held constant, in order to examine the effect on diagnosis and diagnostic ratings (e.g., Becker & Lamb, 1994; Crosby & Sprock, 2004; Ford & Widiger, 1989; Warner, 1978).

Becker and Lamb (1994) asked clinicians (social workers, psychologists, and psychiatrists) to provide a diagnostic evaluation of a case containing equal numbers of symptoms from the DSM-III-R criteria for posttraumatic stress disorder (PTSD) and BPD. Half the participants received a "male" version of the case while the other half received a "female" version. The results showed that the female version of the case was more likely to receive a diagnosis of BPD than the male version, which received more diagnoses of PTSD.

More recently, a study conducted by Boggs et al. (2005), showed that ratings of functional impairment for a number of the BPD criteria were consistently lower for women compared to men. In other words, female patients exhibiting symptoms such as "chronic feelings of emptiness" were seen as less impaired (in social or occupational functioning) than their male counterparts. As the authors pointed out, such a finding suggests that the criteria for BPD may have a gender bias. Notably, none of the criteria for the other three personality disorders examined exhibited gender differences in level of impairment.

Clinician variables in BPD diagnosis. Becker and Lamb (1994) also examined the clinician variables that were associated with a diagnosis of BPD. First, they found that gender of the clinician did not affect diagnosis, with both male and female clinicians

tending to disproportionately assign a diagnosis of BPD to the female patient. However, they did find that both the age of the clinician and the proportion of time spent in direct patient contact had an impact on diagnostic decisions. Specifically, younger clinicians and those who spend the majority of their time directly working with patients tended to assign the diagnosis of BPD more readily than older clinicians and clinicians who spend less of their time in direct practice. These latter findings suggest that professional and demographic characteristics of the clinicians may contribute to a misdiagnosis of BPD in clinical settings and point to the need for further investigation into these areas.

Another variable that was not found by Becker and Lamb (1994) to lead to a higher frequency of using BPD as a diagnosis was the clinicians' profession. Although psychiatrists tended to be less inclined to assign the diagnosis of PTSD compared to other professional groups, there were not significant differences in the diagnosis of BPD. Ochoa and Morey (1990) also found that psychologists and psychiatrists tended to be similar in the manner in which they went about assigning the diagnosis of BPD, with both groups relying heavily on the DSM-III criteria. However, they found that psychologists and psychiatrists differed in the importance they placed on various criteria. Specifically, psychiatrists tended to view symptoms associated with management issues (e.g., suicidal ideation, self-mutilation) as more important when assigning a diagnosis of BPD, whereas psychologists viewed affective instability as the most important determinant.

Morey and Ochoa (1989) also examined clinician variables that were predictive of personality disorder diagnoses, including BPD. However, they utilized a different methodology in which they asked experienced psychologists and psychiatrists to identify a patient they had diagnosed with a personality disorder, provide the patient's diagnosis,

and indicate which personality disorder criteria were present for the patient. They compared the patient's diagnosis based on the symptom ratings (i.e., labeled the "criterion diagnosis") to the assigned diagnosis (i.e., "clinical diagnosis") in order to investigate clinician adherence to diagnostic criteria and factors that predicted over-diagnosis and under-diagnosis of personality disorders. First, they found a high degree of inconsistency between clinical and criterion diagnoses. With respect to a diagnosis of BPD, unlike Becker and Lamb (1994), they found that gender of the clinician predicted over-diagnosis. Female clinicians were more likely to diagnose BPD without evidence to justify the diagnosis (i.e., they did not meet sufficient criteria for BPD), whereas male clinicians were more likely to not diagnose BPD when sufficient criteria were present (i.e., under-diagnosis). However, their results were consistent with Becker and Lamb with regard to clinician experience, with less experienced clinicians over-diagnosing BPD. Other clinician variables predictive of over-diagnosis of BPD were race (i.e., whites over-diagnosed) and theoretical orientation (i.e., dynamically-oriented clinicians over-diagnosed). Moreover, the presence of specific symptoms were associated with over-diagnosis of BPD (i.e., suicide gestures, self-damaging impulsivity, affective instability), whereas other symptoms predicted under-diagnosis of BPD (i.e., indifference to praise or criticism, absence of manipulative suicide threats). The diagnostic features most associated with a diagnosis of BPD were self-destructive behavior, anger, suspiciousness, and interpersonal difficulties. Overall, they concluded that clinicians use their own clinical judgment when assigning personality disorder diagnoses rather than the DSM criteria. Because they studied actual patients seen by clinicians rather than vignettes, their findings may have increased generalizability. On the other hand, they

asked clinicians to identify patients with whom they had a minimum of 10 hours of contact during the past year. Therefore, their results might not generalize to the typical clinical situation where a diagnosis is assigned after the intake interview or an initial assessment. Nevertheless, their findings suggest that clinician variables may influence the diagnosis of BPD and result in diagnostic bias.

Blashfield and Herkov (1994) used the identical methodology as Morey and Ochoa (1989) in an attempt to replicate their findings. They also found considerable disagreement between "clinical" and "criterion" diagnoses; however, they reached a somewhat different conclusion. Rather than error, they interpreted the results as reflecting two different methods of assigning a diagnosis: the DSM approach of carefully evaluating whether a patient meets specific diagnostic criteria and the prototype approach in which patients are compared to a representative example of the diagnosis. According to these authors, neither approach is necessarily superior, although they result in different diagnoses. In terms of predictors of diagnostic bias, Blashfield and Herkov failed to replicate most of the predictors of underdiagnosis of BPD found by Morey and Ochoa (1989) and did not find clinician gender to be a significant predictor of overdiagnosis of BPD. Their results did replicate Morey and Ochoa's (1989) findings that the presence of certain symptoms (e.g., recurrent suicide gestures, self-damaging behaviors, affective instability) is predictive of overdiagnosis of BPD.

Negative emotional reactions to BPD. A wide body of research (e.g., Gallop, Lancee, & Garfinkel, 1989; Lenzenweger & Cicchetti, 2005; Markham, 2003; McIntyre & Schwartz, 1998; Nehls, 1998; Reiser & Levenson, 1984) has shown that individuals diagnosed with BPD elicit strong negative reactions from clinicians and other treatment

staff and that there is considerable stigma associated with this diagnosis. According to Nehls (1998), treatment providers tend to describe patients with BPD in highly derisive terms such as "hateful," "manipulative," and "non-compliant." In an illustration of this phenomenon, Gallop, Lancee, and Garfinkel (1989) conducted a vignette study in which 113 nurses on an inpatient psychiatric unit were presented with a case briefly describing characteristics of a hypothetical patient presenting for treatment. The nurses were told that the patient was in her late 20's, had a history of multiple psychiatric admissions, and had previously reported negative attitudes about treatment. One group of nurses received a vignette indicating that the patient was diagnosed with schizophrenia whereas the other group of nurses was told that the patient was diagnosed with BPD. The nurses were then asked to rate their predicted level of empathy for the patients and write several empathic statements regarding the patient. The results of this study revealed that those nurses who were told that the patient's diagnosis was BPD reported significantly lower levels of empathy than those who were told that the patient had a diagnosis of schizophrenia. Moreover, the patient described as having BPD seemed to elicit more disdainful comments than the patient diagnosed with schizophrenia. Since the nurses in this study were provided with identical descriptions of the patient (with the exception of the diagnosis), these findings suggest that a diagnosis of BPD may result in negative responses from treatment providers.

Recent studies (Cleary, Siegfried, & Walter, 2003; Markham, 2003) have reached similar conclusions, showing that nurses and paraprofessionals reported less optimism about working with BPD patients compared to other groups and described their work as more difficult than working with other patient populations. Interestingly, a study

conducted by Markham and Trower (2003) found one potential explanation for these findings, as nurses were more likely to describe observed pathological behavior as being within the BPD patients' control when compared to patients diagnosed with depression or schizophrenia. Moreover, the ratings of perceived control were negatively correlated with ratings of sympathy towards the patient. Thus, the behaviors exhibited by BPD patients in this study were seen as voluntary and therefore unworthy of sympathy. Such reactions obviously have an impact on the way treatment is provided and shed light on the dangers of an inaccurate diagnosis of BPD.

Additional research shows that such emotional reactions are not restricted to nurses or other support staff on an inpatient unit. A study by Bongar, Markey, and Peterson (1991) surveying psychiatrists on "difficult and dreaded" patients found that the majority of those identified as such by psychiatrists had a diagnosis of BPD. More recently, McIntyre and Schwartz (1998) asked 155 licensed psychotherapists (professional degrees not specified) to listen to audiotape recordings of an intake session with a patient diagnosed with either Major Depression or BPD (but they were not provided with information regarding the diagnosis). After listening to the recording, participants were asked to rate their reaction to the patient on two measures: The Impact Message Inventory (IMI) (Keisler, 1993) and the Stress Appraisal Scale (SAS) (Carpenter & Suhr, 1998). As will be discussed in more detail later in this paper, the IMI is a 90-item measure of personal reactions to other individuals and has been extensively used in the study of therapist reactions towards patients. In contrast, the SAS is a 36-item measure of therapist attitudes about entering into a therapeutic relationship with a client prior to beginning treatment. Therapists who listened to the recording of a patient

diagnosed with Major Depression rated the patient as being significantly more friendly and submissive (i.e., willing to accept responsibility, deferential to the therapist), and reported more feelings of caring when compared to the BPD case. In contrast, the therapists who listened to the patient diagnosed with BPD rated the patient as being more dominant and reported greater feelings of hostility and emotional detachment. Moreover, these latter therapists also reported a belief that therapy with the patient diagnosed with BPD would likely be unproductive. A significant effect was also found in this study for years of clinical experience. Specifically, clinicians with fewer years of clinical experience reported more negative emotional reactions than clinicians with more experience (McIntyre & Schwartz, 1998).

Another study (Hillman, Stricker, & Zweig, 1997) examining the effect of patient age on the diagnosis of personality disorders found that psychologists showed significantly more negative reactions to a patient with Major Depressive Disorder with BPD compared to a patient with only Major Depressive Disorder. The patient with comorbid BPD was rated as significantly lower in functioning, having more severe symptoms, a poorer outcome, and being less ideal for their practice. Several unpublished dissertations have also demonstrated that psychologists have negative attitude towards patients with BPD (i.e., Brigg-Phillips, 2004; Giacalone, 1997).

The above findings suggest that the attitudes of clinicians towards patients exhibiting symptoms of BPD are markedly distinct from their attitudes towards other patient groups. In other words, patients displaying "borderline" features seem to elicit negative emotional reactions from caregivers that have the potential to be detrimental to the therapy process. Thus, it is possible that these negative emotional reactions may have

a similar detrimental impact on diagnostic assessment and may be an additional, previously unrecognized, form of diagnostic bias. Moreover, it appears that inexperienced clinicians may be especially susceptible to this form of bias given that they report more intense emotional reactions to BPD symptomology (McIntyre & Schwartz, 1998) and have demonstrated higher rates of using the BPD diagnosis compared to more experienced clinicians (Becker & Lamb, 1994; Morey & Ochoa, 1989). As Becker and Lamb suggested, BPD is quickly becoming a diagnosis that is not only used frequently during times of uncertainty but is used as a form of denigration. Perhaps, then, clinicians are using the diagnosis of BPD as a way to convey their negative reactions towards patients in an indirect manner. Since there are no empirical studies directly examining the effect of NER on diagnosis, it is necessary to turn to other sources in order to obtain a more comprehensive understanding of the diagnostic process.

Decision-Making Research

Research on the process of human decision-making has been applied to a wide variety of areas, including economics, the legal system, the military, and medicine (Maule & Svenson, 1993). Broadly defined, research on decision-making attempts to determine the specific factors of an object or situation that are influential in determining the judgments of an individual. Methodologically, decision-making research often presents individuals with a specific goal, such as purchasing the best car, and provides them with several choices, each of which have different attributes. According to Maule and Svenson (1995), decision-making methodologies differ in terms of whether they emphasize structural or process analysis. In a structural analysis approach, the participant is presented with all of the attributes of each object. Thus, the participant's ultimate

decision is guided by determining which object possesses the most attractive attributes based on the information provided. In contrast, process approaches to decision-making are based on the belief that in the real world individuals are rarely provided with all of the attributes for a particular object. As a result, process approaches study decision-making at the level of attribute selection. Studies that emphasize a process approach typically provide participants with a list of attributes to choose from in order to determine those that are the most crucial to the decision-making. Maule and Svenson (1995) argued that process based approaches provide researchers with a clearer picture of the various decision-making strategies employed by individuals (e.g., risk-avoidant) than do structural approaches. Moreover, process approaches have yielded interesting findings into the powerful role that environmental factors (e.g., time-pressure and stress) can play on the process of attribute selection.

Perhaps more than any other decision-making task, process approaches appear to be particularly applicable to the study of medical decision-making. In such settings, patients often present with vague complaints and it becomes the task of clinician to form a diagnosis based on this and other information solicited during an interview. In order to reach a decision, the clinician must decide which information is appropriate to solicit and which should be ignored. However, a number of studies have shown that clinicians vary considerably in the strategy they use to select this information due to personal characteristics (e.g., experience level) and environmental circumstances (e.g., time pressure or stress). In a review of this issue conducted by Gruppen, Wolf, and Billi (1991), the authors reported that the process of clinical decision-making can be regarded as the testing of various hypotheses. When formulating a diagnosis under this model,

referred to as the hypothetico-deductive model, clinicians develop a limited number of hypotheses early in their encounter with the patient. These hypotheses then guide their selection of information that is gathered from the patient so that the information gathered either confirms or disconfirms these hypotheses. Once the information is gathered, the clinician then integrates the information and chooses the hypothesis that has the most support based on the evidence. Although most medical decision-making tasks fit neatly into the hypothetico-deductive model (Bergus & Hamm, 1995), this strategy is laden with imperfections. For instance, each step in this model is dependent on the successful completion of the previous one (Gruppen, Wolf, & Billi, 1991). As a result, if a clinician selects a set of incorrect hypotheses in the beginning, this will obviously influence the content of information that is solicited from the patient. Even if the initial hypotheses that are selected are correct, the ultimate selection of a diagnosis is largely dependent on one's expertise in soliciting appropriate information.

Heuristics and prototype matching. Other approaches to clinical decision making, particularly focusing on those who have high levels of expertise, have regarded the process of selecting a diagnosis as one of pattern recognition (Elstein & Schwarz, 2002). The speed at which experienced clinicians arrive at a diagnosis suggests that rather than developing and testing hypotheses, more experienced clinicians tend to use a pattern recognition approach in which the symptoms are matched to either representative instances (e.g., exemplars) or mental models (i.e., prototypes) of the diagnosis (Elstein & Schwarz, 2002). Although this strategy can be expeditious, its success largely depends on the appropriateness of the forms being retrieved (Elstein & Schwarz, 2002). In addition, the hypothesis testing approach is still used by experienced clinicians when

faced with difficult cases. In both the hypothetico-deductive and pattern matching approaches to decision making, the clinician is taking advantage of a heuristic. Simply defined, a heuristic is a cognitive strategy that is employed in order to make complex tasks simpler (Hall, 2002). Representativeness heuristics can help prevent clinicians from having to needlessly gather information that is irrelevant to the task at hand and form the most accurate decision in the shortest time possible.

However, the use of heuristics may also result in decision making errors such as neglect of base rates, overestimation of the likelihood of vivid or easily recalled events, and underestimation of the frequency of more ordinary or difficult to recall instances. In addition, there are a number of other cognitive biases that can lead to errors in decision making, such as representativeness (i.e., overestimation of the probability of a case by its similarities to the category or prototype), probability transformations (overweighing of small probability events and underweighing of large probabilities), order effects (information presented later in a case is given more weight), and confirmation biases in which there is emphasis on findings consistent with a hypothesis and minimization of disconfirming information (Elstein & Schwarz, 2002).

Hall (2002) observed that heuristics are more likely to be used under conditions of uncertainty; however, heuristics and prototypes frequently cause clinicians to commit basic procedural errors. For instance, clinicians may lose sight of base rates of certain diagnoses when making quick decisions and may overestimate or underestimate the actual frequency of a diagnosis in the population. Secondly, clinicians may become insensitive to the reliability of information provided to them and fail to question the source of the information. Finally, clinicians may exhibit a tendency to over-weigh the

importance of diagnostically irrelevant information to their ultimate decision when using a heuristic. Thus, it appears that the accuracy of a diagnostic decision may be inversely related to the clinician's reliance on a heuristic.

Blashfield and Herkov (1996) suggested that clinicians frequently use a prototype matching approach to diagnose personality disorders, and that it is as valid an approach as the use of specific diagnostic criteria outlined in the DSM-IV. In fact, Westen (1997) has argued that the pattern matching (i.e., prototype) approach may be especially useful in the case of diagnosing personality disorders, and that it is actually more valid and consistent with clinical practice than comparing a patient's symptoms to a list of symptoms as in the DSM-IV. In Westen's prototype-matching approach, clinicians rate the degree of match between their client and clinician-derived prototypes, and a diagnosis is assigned if the degree of match reaches a pre-determined threshold. Although this process may more accurately mirror the decision-making process of experienced clinicians, the decision-making literature suggests that this approach may be more susceptible to bias (e.g., Hall, 2002). Moreover, Westen has suggested that when diagnosing personality disorders, clinicians place increased importance on the manner in which the patient interacts with them and the emotional responses that the patient's behaviors elicit, rather than specific diagnostic criteria. Therefore, clinicians may be especially prone to being influenced by their emotional responses, rather than the diagnostic criteria, when diagnosing BPD and other personality disorders.

Decision-Making Under Stress. In examining the decision-making research, there is some evidence that emotions and stress may result in decreased performance and influence decision-making biases and use of heuristics. Under stress, specifically in the

form of time pressure, the use of heuristics to guide decision-making seems to increase (Edland & Svenson, 1993). In a review of the literature on this topic, Edland and Svenson noted that time pressure has been shown to cause individuals to engage in more risk-avoidant decision-making. In other words, individuals under time pressure tend to select fewer attributes of their alternative choices and focus only on those that could lead to negative consequences. By doing so, individuals under time pressure may ignore potentially relevant attributes and therefore make more faulty decisions. Overall, though, research on the effect of time pressure has shown that individuals tend to narrow their focus to those attributes that are viewed as the most important for a particular decision.

For instance, Edland (1993) conducted a study on the decision-making habits of students asked to rate the academic potential of a candidate for a school psychology program. The results of this study showed that those in the time pressure condition tended to base their decision almost entirely on what they viewed as the most important attribute (i.e., psychology grades), whereas those in the no time pressure group tended to give more weight to what they viewed as less important attributes (i.e., other grades). This phenomenon of tunnel vision, which has been consistently observed in the literature on time pressure and decision-making, has been shown to have negative consequences for the accuracy of decisions. For instance, a review of literature on the effects of emergency situations on aviators conducted by Orasanu (1997) reported that increased stress often led to aviators failing to assess situations thoroughly before making a decision, and this effect was more salient for complex emergency situations (e.g., misleading cues). According to Svenson (1997), the reasons for changes in decision-making strategies under stress may be due to a decreased allocation of cognitive

resources for the decision making task. If more resources are being used to cope with the stress (e.g., maintaining composure), then fewer resources would be available to engage in more adaptive decision-making strategies. Another explanation may be the tendency to overestimate risk as described by Hall (2002) for medical decision making.

In general, affect, both positive and negative, has been found to affect decision making, but the relationship may be complex, and may be different for stable versus more transient affect (Jundt & Hines, 2002). Although it may be counterintuitive, it has generally been suggested that high levels of negative affect may result in more accurate, in-depth, and systematic thinking, whereas high levels of positive affect results in increased reliance on heuristics. However, negative affect may lead to more biased decision making on personally relevant ratings (Jundt & Hines, 2002).

Hall (2002) pointed out that the uncertainty of clinical decision making itself creates stress for both students and seasoned professionals that can contribute to diagnostic errors. Cumming and Harris (2001) investigated the role of anxiety in clinical decision making using a sample of advanced undergraduate medical radiation students. They noted that clinical decision making may be particularly anxiety-provoking for early career professionals, but they were also interested in the effect of trait anxiety on clinical decision making. They examined both naturalistic (scores on an anxiety inventory) and experimentally-induced anxiety and found that both negatively affected decision making for some tasks, with anxiety associated with increased use of heuristics and simpler decision-making strategies. In contrast to their prediction, anxiety had an effect on the primary (i.e., simpler) decision-making task but had little effect on the more challenging task. They hypothesized that anxiety may have caused participants to redirect their

attention from the primary task causing a decline in performance, but had less effect on the secondary task because it did not engage their attention (i.e., there was less of a competition for cognitive resources).

Diagnosis of BPD

In the case of clinicians diagnosing BPD, it can be argued from the literature that clinicians experience considerable negative emotional reactions when interacting with patients with BPD features relative to other patients (i.e., Gallop et al., 1989; Lenzenweger & Cicchetti, 2005; Markham, 2003; McIntyre & Schwartz, 1998; Nehls, 1998; Reiser & Levenson, 1984). Given that stress, anxiety, and negative affect have been shown to have an impact on the process of decision-making (i.e., Cumming & Harris, 2001; Edland & Svenson, 1993; Orasanu, 1997), clinicians interviewing patients with "borderline" features (e.g., anger toward the clinician, suicidal threats) may increase their use of compensatory heuristics. In addition, the heterogeneity of the symptoms and vagueness of the diagnostic criteria for BPD (i.e., Gunderson et al., 1991) may increase uncertainty, which also increases the likelihood of using heuristics and the possibility of biases affecting decisions (i.e., Hall, 2002). For instance, clinicians may cease to develop alternative hypotheses and narrow their focus in such situations, thereby focusing only on information consistent with a BPD diagnosis and ignoring other information. They may also overestimate the importance of vivid and easily recalled information, so that a few BPD symptoms may be enough to trigger a diagnosis of BPD (i.e., Elstein & Schwarz, 2002). In addition, clinicians who use a prototype approach to diagnosis may be more susceptible to diagnostic bias (i.e., Hall, 2002). Moreover, less experienced clinicians, due to their stronger emotional reactions towards patients with BPD features (i.e.,

McIntyre & Schwartz, 1998), may be more likely engage in such strategies, which could explain the higher rates of the BPD diagnosis assigned by those clinicians in previous studies (i.e., Becker & Lamb, 1994; Morey & Ochoa, 1989).

Current Investigation

The present study examined the notion that psychologists' negative emotional reactions (NER) towards behaviors and symptoms associated with BPD might provide further understanding of the over-diagnosis of BPD. Clinicians were asked to read two case vignettes describing female patients who exhibited two criteria for BPD along with behavior consistent with Nehls' (1998) description of behaviors seen with BPD patients (e.g., that they are difficult, angry, noncompliant, hateful) that elicited negative reactions among professionals. Each vignette contained two different criteria for BPD since it was thought that different BPD criteria and behaviors would affect clinician NER and BPD diagnoses. Because each case contained only two symptoms of BPD, both cases were subthreshold for the diagnosis. A female patient was used in the case vignettes because the majority of patients diagnosed with BPD are women (APA, 1994) and clinicians tend to associate this disorder with female patients (Becker & Lamb, 1994; Morey & Ochoa, 1989). The order of the cases was counterbalanced.

For each case, clinicians were asked to: rate the applicability of a selected set of DSM-IV Axis I and II diagnoses to the case; assign an Axis I and Axis II diagnosis and rate their confidence in their diagnoses; rate severity, prognosis and likelihood that the patient would benefit from treatment; and rate the applicability of a series of DSM-IV Axis I and II criteria (including all of the BPD criteria) to the case. The symptom ratings allowed for a comparison of the symptoms they see as present in the case and the

diagnosis they assigned, paralleling the methodology used by Morey and Ochoa (1989) and Blashfield and Herkov (1996). The order of the diagnostic and symptom ratings was counterbalanced to control for order effects. In addition, these two orders were used to simulate the two decision-making strategies noted by Blashfield and Herkov (1996). Participants asked to assign a diagnosis before evaluating symptoms were being asked to use a prototype approach, whereas those asked to rate the presence of DSM-IV symptoms prior to assigning a diagnosis were being asked to use a DSM-IV approach. It has been suggested that clinicians who utilize the prototype approach may be more prone to bias (i.e., Hall, 2002) and to over-diagnosing personality disorders relative to their symptom ratings (i.e., Crosby & Sprock, 2004). After the diagnostic and symptom ratings were assigned, clinicians were asked to rate their degree of NER towards the patient using a series of items selected from the Impact Message Inventory (IMI) (Keisler, 1993). Finally, participants were asked to provide demographic information and information about their professional experience. The entire study was conducted on the Internet.

The following hypotheses were proposed:

1. Compared to clinicians reporting lower levels of NER, clinicians reporting higher levels of NER in response to the cases will see the cases as more representative of BPD and more pathological as evidenced by:
    a. Assigning higher ratings of the applicability of BPD to the cases.
    b. Diagnosing BPD more frequently.
    c. Rating the pathology of the case as more severe, the prognosis poorer, and the patient as less likely to benefit from treatment.
2. The effects reported in hypothesis #1 will be moderated by clinician variables

including clinicians' years of experience and percent of time spent in direct patient contact. Due to the variable findings regarding clinician gender, no hypothesis concerning clinician gender was made. The specific hypotheses were that:

a. Compared to clinicians with more years of experience, clinicians with fewer years of experience will report higher levels of NER, and as a result: assign higher BPD ratings, diagnose BPD more frequently, rate the case as more severe, the prognosis lower, and the patient as less likely to benefit from treatment.

b. Compared to clinicians with less frequent patient contact, clinicians with more direct patient contact will report higher levels of NER, and as a result: assign higher BPD ratings, diagnose BPD more frequently, rate the case as more severe, the prognosis as poorer, and the patient as less likely to benefit from treatment.

3. The order of symptom and diagnostic ratings will have an effect on the diagnostic ratings such that:

a. Those who are asked to assign a diagnosis first and then rate symptoms (i.e., prototype approach) will assign higher ratings of the applicability of BPD to the case and diagnose BPD more frequently.

b. Those who are asked to rate symptoms first and then assign a diagnosis (i.e., DSM-IV approach) will assign lower ratings of the applicability of BPD to the case and diagnose BPD less frequently.

4. When participants' diagnoses are compared with their symptom ratings, there will be within-subject inconsistency and biases such that:
    a. There will be considerable disagreement between clinicians' "criterion diagnoses" of BPD, determined by their BPD symptom ratings, and assigned diagnoses of BPD suggesting over-diagnostic or under-diagnostic bias.
    b. Clinicians reporting higher levels of NER in response to the cases will be more likely to overdiagnose BPD (i.e., diagnose BPD in the absence of endorsing a sufficient number of DSM-IV symptoms for the diagnosis) than those with lower levels of NER.
    c. Those asked to assign a diagnosis first and then rate symptoms (i.e., prototype approach) will be more likely to overdiagnose BPD (i.e., diagnose BPD in the absence of endorsing a sufficient number of DSM-IV symptoms for the diagnosis) than those asked to rate symptoms first and then assign a diagnosis (i.e., DSM-IV approach).

Chapter 2

METHOD

Design

In order to test the primary hypothesis that clinicians' negative emotional reactions (NER) will detrimentally influence the quality of diagnostic decision-making, a correlational design was used with two primary predictors (IMI Dominant and Hostile scales). The primary criterion variables were BPD diagnosis and BPD representativeness ratings. Secondary predictors included clinician years of experience and percent of direct patient contact. Secondary criterion variables included: ratings of severity, prognosis, and likelihood of benefiting from treatment; and within-subject overdiagnosis and underdiagnosis (i.e., a comparison of "criterion diagnoses" based on their symptom ratings to assigned diagnoses). There was also an experimental component in which individual participants were randomly assigned to receive the diagnostic ratings before the symptoms ratings (i.e., prototype approach) or the symptoms ratings before the diagnostic ratings (i.e., DSM-IV approach) to simulate the two approaches to diagnostic decision making (i.e., Blashfield & Herkov, 1994) and to examine their effect on diagnostic bias. Two case vignettes were used, each with two different symptoms of BPD, to control for the possibility that the unique aspects of a particular case vignette

could account for the findings, and each case was analyzed separately. Order of the cases was counterbalanced, with participants randomly assigned to one of the two orders.

The primary hypothesis was tested by examining the relationship of the NER ratings with the assigned diagnosis, diagnostic ratings, and ratings of severity, prognosis, and likelihood of benefiting from treatment. Additional analyses examined the relationship of the two clinician variables (i.e., years of clinical experience, percent time spent in direct patient contact) with NER, BPD diagnosis, and BPD diagnostic ratings, and the effect of the two experimental conditions (i.e., prototype approach vs. DSM-IV approach) on the dependent (criterion) variables. Finally, the consistency between the criterion diagnosis and the assigned diagnosis was used to determine intra-subject bias, and the relationship of the predictors to over- and under-diagnostic bias was examined.

Participants

A power analysis suggested that 67 participants were needed to detect a medium effect size with a power of 80 and alpha of .05 for multiple regression using two predictors (Cohen, 1992). With the additional three secondary predictors (the two clinician variables plus order of diagnosis and symptom ratings), a minimum of 91 participants would be needed. Based on response rates of 10-20% found in similar studies (e.g., Crosby & Sprock, 2004), a random sample of 1000 names was obtained from the online APA directory of psychologists using a table of random numbers in a ratio representative of the number of psychologists in each state. The email address for each participant was obtained from the listing in the online directory.

After an initial email invitation and one reminder email that was sent to the selected 1000 participants, a total of 90 psychologists participated in the study; however,

approximately 1/3 of the emails sent out were returned as invalid, bringing the response rate to about 13%. In order to ensure a sufficient number of participants, an additional sample of 200 psychologists was randomly selected and invited in a manner identical to the initial sample, resulting in an additional 12 responses, for a total of 102 participants. However, one participant was excluded for large amounts of missing data, one participant was excluded because he identified himself as retired, and two were excluded because they were found to be duplicate entries in the data file. This brought the final number of participants to 98, which was deemed adequate for the analyses.

Tables 1 and 2 present the demographic and professional characteristics of the participants. Participants had a mean age of about 51 years old and an average of approximately 18 years of clinical experience. There were slightly more women than men, and most participants identified themselves as "Caucasian." All had doctorates, the majority a Ph.D., and clinical psychology was the most frequent area of study. On average, the participants reported that they spent more than half of their time in direct patient contact, with the majority of this time being with adults ages 18-64. The most commonly identified theoretical orientation was "eclectic" followed closely by "cognitive-behavioral." "Private practice" was the most frequently reported employment setting, although a variety of settings were represented.

Materials

All of the materials for this study were located on an Internet website established by the researcher. The website contained the following materials: the case vignettes, the diagnostic and symptom ratings, the Impact Message Inventory items, and the demographic questionnaire.

Table 1

Characteristics of Participants: Means and Standard Deviations

	Mean	(SD)
Age	51.39	(9.79)
Year of Doctorate	1986.46	(9.61)
Years Clinical Experience	18.08	(9.86)
DSM-IV Experience[a]	5.22	(1.49)
Percent Time Direct Contact[b]	55.90	(33.71)
Percent Time Children[c]	15.00	(23.75)
Percent Time Adolescents[c]	17.75	(23.01)
Percent Time Young Adults[c]	17.26	(17.92)
Percent Time Adults[c]	26.31	(20.51)
Percent Time Middle Age[c]	16.66	(14.58)
Percent Time Older Adults[c]	5.32	(10.30)

[a] Rated on a scale of 1 (none) to 7 (very experienced)
[b] Percent of time spent in direct patient contact
[c] Percent of time working with specific age groups

Table 2

Characteristics of Participants: Frequencies and Percentages

	Freq (%)
Gender	
Male	46 (46.9%)
Female	52 (53.1%)
Ethnicity	
Caucasian	93 (94.9%)
African American	2 (2.0%)
Hispanic	1 (1.0%)
Asian	1 (1.0%)
Not Reported	1 (1.0%)
Type of Degree	
Ph.D.	77 (78.6%)
Psy.D.	17 (17.3%)
Ed.D.	3 (3.1%)
Not Reported	1 (1.0%)
Field of Degree	
Clinical Psychology	82 (83.7%)
Counseling Psychology	10 (10.2%)
School Psychology	1 (1.0%)
Education	1 (1.0%)

Table 2 (*Continued*)

	Freq (%)
Other	3 (3.1%)
Theoretical Orientation	
Eclectic	35 (35.7%)
Cognitive-Behavioral	30 (30.6%)
Psychodynamic	24 (24.5%)
Humanistic	5 (5.1%)
Other	4 (4.1%)
Primary Setting of Practice	
Private Practice	44 (44.9%)
Other	23 (25.5%)
Academic Department	10 (10.2%)
Community Mental Health	6 (6.1%)
Private Psychiatric Facility	5 (5.1%)
General Medical Hospital	4 (4.1%)
University Medical School	2 (2.0%)
State Psychiatric Facility	2 (2.0%)
Correctional Facility	2 (2.0%)

Case vignettes. Each participant was presented with two case vignettes (see Appendix B) developed for this study that described a female patient who exhibited two of the nine DSM-IV criteria for BPD. The case vignettes contained two different

symptoms of BPD because it was expected that the presence of specific symptoms might influence the results. Given the higher prevalence rates for BPD among women (APA, 1994) and clinicians' tendency to assign the diagnosis more often to women (e.g., Becker & Lamb, 1994), a female client was used in the vignettes. Both vignettes were limited to approximately 500 words in order to fit entirely on a single screen on the computer, and order of the cases was counterbalanced.

Case One contained the following two symptoms of BPD: "a pattern of unstable and intense interpersonal relationships characterized by alternating between extremes of idealization and devaluation" and "inappropriate, intense anger or difficulty controlling anger (e.g., frequent displays of temper, constant anger, recurrent physical fights)." Case Two contained the following symptoms of BPD: "frantic efforts to avoid real or imagined abandonment" and "affective instability due to a marked reactivity of mood (e.g., intense episodic dysphoria, irritability, or anxiety usually lasting a few hours and only rarely more than a few days)." Because a minimum of five of the nine criteria are necessary to receive the diagnosis (see Appendix A), both cases were subthreshold for BPD. The selection of the criteria was based on symptoms of BPD thought to be most likely to produce NER (e.g., inappropriate, intense anger) as well as those found to be strongly associated with a diagnosis of BPD (e.g., Ochoa & Morey, 1990). The case vignettes also contained behaviors cited in the Nehls (1998) review as prompting negative attitudes towards patients diagnosed with BPD (i.e., resistant to and noncompliant with treatment, animosity directed towards the treatment staff). Although both cases were intended to elicit NER from participants, it was expected that the vignettes would produce varying levels of NER throughout the sample. Pilot testing was

conducted in order to determine if the cases contained the intended BPD symptoms and to test the effectiveness of the cases in eliciting varying levels of NER. A description of the pilot studies is presented in Appendix C.

Diagnostic and symptom ratings. Participants were asked to rate the applicability of a series of diagnoses to the cases using a 7 point scale (1 = not at all representative to 7 = very representative). The diagnoses included each of the DSM-IV personality disorders (plus the two in the DSM-IV appendix) as well as Axis I diagnoses relevant to the information presented in the cases (i.e., Major Depressive Disorder, Adjustment Disorder, Dysthymic Disorder, Bipolar I and II Disorders, Intermittent Explosive Disorder). In total, the participants were asked to rate the applicability of 18 diagnoses. Participants were then asked to select the one best Axis I diagnosis and Axis II diagnosis, and rate their level of confidence in each diagnosis using a 7 point scale (1 = not at all confident to 7 = very confident). Participants were also asked to rate severity, prognosis, and the likelihood of the patient benefiting from treatment, also using a seven-point scale, and to rate the applicability of a series of 42 symptoms for the case. The symptoms consisted of a subset of the DSM-IV criteria for each of the diagnoses presented to them, and included all nine DSM-IV criteria for BPD. The purpose of the symptom ratings was to determine which symptoms of BPD the rater viewed as present in the case in order to obtain a "criterion diagnosis" with respect to BPD (i.e., Blashfield & Herkov, 1996; Morey & Ochoa, 1989; Sprock & Crosby, 2004). The "criterion diagnosis" could then be compared with the assigned diagnosis to determine within-subject over-diagnostic and under-diagnostic bias for BPD. The order of presentation of the diagnostic and symptom ratings was counterbalanced. See Appendix D for the diagnostic and symptom ratings.

Impact Message Inventory (IMI; Keisler, 1993). After the diagnostic and symptom ratings were completed, participants were asked to provide ratings of their negative emotional reactions towards the case using a subset of the items from the Impact Message Inventory (IMI; Keisler, 1993). The IMI was selected to measure this variable as it has been used extensively in previous studies to measure the impact that psychiatric patients have on their therapists (e.g., McIntyre & Schwartz, 1998; Simon, Gaul, Freidlander, & Heatherington, 1992; Stephens, Hokanson, & Welker, 1987; Swaney & Stone, 1990).

The IMI is a 90-item inventory that yields 15 subscales reflecting varying types of interpersonal impacts. The inventory asks participants to rate the similarity of a statement to his/her own personal reaction (e.g., "When I am with this person, he or she makes me feel annoyed") on a 1 to 4 scale. Given that this inventory has primarily been used to assess the impact of actual personal encounters (as opposed to vignettes), the items were re-worded to be more consistent with the current study (e.g., "When I am with this person she makes me feel annoyed"), and participants were instructed to imagine that they were interacting with the patients in the cases. Although the IMI yields 15 subscales, the scale's author (Keisler, 1993) recommends that these subscales be grouped into four different clusters (i.e., Dominant, Hostile, Submissive, and Friendly) due to high intercorrelations between the subscales. The four clusters are reported to have adequate reliability, with alpha coefficients ranging from .84 to .89. Given McIntyre and Schwartz's (1998) finding that therapists typically report elevated scores on the Dominant and Hostile clusters in response to BPD patients, only the items from these two clusters were used. A total of 24 items comprise the Dominant and Hostile clusters (i.e.,

scales) of this measure and are listed in Appendix E. Also, Appendix F contains the author's full description of the characteristics associated with both scales.

Demographic questionnaire. Finally, each participant was asked to provide demographic information (e.g., gender, age, ethnicity) and information about his or her professional training and experience (e.g., type of degree, years of experience, percent of time in direct patient contact). The specific items for the demographic questionnaire are listed in Appendix G.

Procedure

The sample was obtained by randomly selecting 1000 psychologists from the APA directory using a table of random numbers, and the email address of each psychologist was obtained from the online APA member directory. Each psychologist was then sent an email inviting them to participate in the study (see Appendix H). The email contained a brief description of the study, a statement ensuring participants that their participation is voluntary, and a password to access the study to ensure that only psychologists in the selected sample could participate. The email also included a link to one of the study's four websites which contained the two orders of symptom and diagnostic ratings and the two orders of cases, and participants were randomly assigned to one of the four websites. In order to increase participation in the study, each participant was offered the opportunity to participate in a raffle to win one of three $50 gift certificates to Amazon.com. Approximately two weeks after receiving the initial invitation, participants were sent an email reminding them to participate in the study. As discussed previously, an additional 200 psychologists were selected and invited to participate to ensure an adequate sample size.

When participants signed onto the website, they received an introductory page informing them that the study was investigating clinical judgments made during intake interviews. Participants were reminded that their participation is voluntary and that they could withdraw from participation at any time. They were told that logging onto the web site, entering the password, and submitting their responses functioned as a statement of understanding of the information provided about the study and participant rights and giving their consent to participate.

After the participants entered their password and logged onto the web site, they were presented with a page containing one of the case vignettes followed by the rating items. The participants were then presented with the second case vignette and instructed to follow the same procedures. After completion of the ratings for the two cases, the participants were presented with the demographic and professional experience questions. At the end of the study, they clicked on the submit button and the data were automatically entered into a database. A thank you page then appeared and participants were invited to provide comments about the study. A link was also presented in which participants could send an email in order to provide their name and address if they were interested in receiving the results of the study or entering their name in the drawing to get one of the three $50 gift certificates.

Chapter 3

RESULTS

Overview of Analyses

All data analyses were conducted using the SPSS Version 11.5 computer software package. The results for Case One and Case Two were analyzed independently and are presented separately. First, descriptive statistics are presented for the dependent measures. The frequencies of assigned diagnoses are presented along with diagnostic confidence ratings and interrater reliability using *kappa* (Cohen, 1960). Next, descriptive statistics are presented for: diagnostic representativeness ratings; symptom ratings; ratings of severity, prognosis, and the likelihood of the patient benefiting from treatment; individual IMI items, and average IMI scale scores. Inter-rater agreement for the dimensional ratings was examined using intraclass correlation (ICC).

A second set of analyses examined the effect of participant characteristics (i.e., years of clinical experience, percent of time in direct patient contact, gender) on the key variables (i.e., diagnoses; diagnostic and symptom ratings; IMI scores; ratings of severity, prognosis, and likelihood of benefiting from treatment). Between-group (chi-square, ANOVA) and correlational analyses were conducted.

A third set of analyses examined the effects of the order of the materials on the participant responses. Specifically, results were compared for participants who were

asked to provide diagnostic ratings first and then symptom ratings (i.e., prototype approach) with those participants who were asked to provide symptom ratings and then diagnostic ratings (i.e., DSM-IV approach). The results were also compared for participants who received Case One and then Case Two and those who received the cases in the opposite order. Chi-square analyses were used to determine if differences in order of ratings or case order significantly affected the diagnoses, and ANOVA's were used to examine the effect of order on the diagnostic representativeness ratings, symptom ratings, IMI scale scores, and ratings of severity, prognosis, and the likelihood of the patient benefiting from treatment.

The next set of analyses examined the relationship between scores on the two IMI scales and the key variables in order to address the primary hypotheses. Initially, a median split was used to separate participants into "high" and "low" groups on each of the two IMI scales, and between-group analyses (chi-square, ANOVA) were used to compare the two groups on the dependent variables. However, given the problems associated with the median split technique (e.g., MacCallum et al., 2002), correlational analyses were also conducted. Specifically, logistic regression was used to explore the degree to which the IMI scales were predictive of a diagnosis of BPD, and multiple regression was used to examine the degree to which IMI scale scores were predictive of BPD representativeness ratings, severity, prognosis, and likelihood of benefiting from treatment.

The last set of analyses examined within-subject diagnostic bias by comparing each participant's symptom ratings to their assigned diagnosis (e.g., Blashfield & Herkov, 1996; Morey & Ochoa, 1989; Sprock & Crosby, 2004). A "criterion diagnosis"

with regard to BPD was determined by examining the number of BPD symptoms rated as present in the case. A rating of four or higher on the seven point scale was used as an indicator that the symptom was present. Assignment of a diagnosis of BPD without endorsing a sufficient number of BPD symptoms (i.e., 5 or more), was considered over-diagnostic bias, whereas failure to assign a diagnosis of BPD but endorsement of five or more symptoms was considered under-diagnostic bias. A logistic regression analysis was used to explore whether IMI scales scores were predictive of over-diagnostic and under-diagnostic bias. In addition, the effect of order of diagnostic and symptom ratings on within-subject diagnostic bias was examined.

Case One

Diagnosis and Confidence Ratings

Table 3 presents the rates at which diagnoses were assigned for Axis I and II. For Axis I, Intermittent Explosive Disorder was the most frequently assigned diagnosis, with 63.3% of the participants assigning it as the Axis I diagnosis that was most representative of the case. As expected, BPD was the most frequent diagnosis assigned for Axis II, with 54.1% of the respondents assigning it as the Axis II diagnosis most representative of the case. Overall, participants reported having a moderate to high level of confidence in their Axis I diagnosis ($M = 3.50$, $SD = 1.68$) and their Axis II diagnosis ($M = 4.50$, $SD = 1.81$).

Reliability. Inter-rater agreement of the Axis I and II diagnoses assigned to the case was examined using the *kappa* statistic (Cohen, 1960). *Kappa* values for Axis I and Axis II diagnoses indicated a moderate level of agreement (see Table 3).

Table 3

Axis I and II Diagnoses Assigned to Case One, with Inter-rater Agreement (Kappa)

Axis I Diagnosis	Freq (%)	Axis II Diagnosis	Freq (%)
Intermittent Explosive Disorder	62 (63.3)	Borderline PD	53 (54.1)
Bipolar II Disorder	11 (11.2)	Narcissistic PD	24 (24.5)
Dysthymic Disorder	11 (11.2)	Histrionic PD	5 (5.1)
Bipolar I Disorder	7 (7.1)	PAPD	4 (4.1)
Adjustment Disorder	3 (3.1)	Antisocial PD	3 (3.1)
Major Depressive Disorder	3 (3.1)	Dependent PD	3 (3.1)
No Diagnosis	1 (1.0)	Paranoid PD	3 (3.1)
		No Diagnosis	2 (2.0)
		OCPD	1 (1.0)
Kappa	*.56*		*.50*

Note. "No Diagnosis" includes participants who either skipped the question in error or did not want to assign a diagnosis to the case. Diagnoses that were not assigned by any participants are not included in the table. PAPD = Passive Aggressive Personality Disorder. OCPD = Obsessive Compulsive Personality Disorder.

Diagnostic Representativeness Ratings

Participants also rated the representativeness of a series of Axis I and II diagnoses for the case. Table 4 lists the average representativeness rating given to each diagnosis

by the participants. For Axis I, Intermittent Explosive Disorder received the highest representativeness rating, in the moderate to high range, with other Axis I diagnoses rated low in representativeness for the case. As expected, BPD received the highest average representativeness rating on Axis II, rated moderate to high in representativeness for the case, followed by Narcissistic Personality Disorder.

Reliability. Interrater agreement of the representativeness ratings was calculated using intraclass correlation (ICC). The raters were treated as variables and ratings as cases. The ICC provides an estimate of the interrater reliability based on the proportion of variance associated with differences between the measured ratings and variability among the raters (Nichols, 1998). Overall, results suggested a moderate level of interrater reliability for the representativeness ratings for Axis I and for Axis II diagnoses (see Table 4).

Symptom Ratings

Table I.1 (Appendix I) lists the average symptom ratings for all nine symptoms of BPD as well as additional highly rated symptoms (i.e., mean rating of 4.0 or higher). Three of the nine symptoms of BPD received a mean rating greater than 4.0, with the most highly rated symptoms being the two intended symptoms for the case: "inappropriate, intense anger or difficulty controlling anger" and "a pattern of unstable and intense interpersonal relationships . . .". The remaining symptoms of BPD, which were not intended to be present in the case, had means of less than 3.0. In addition to the BPD symptoms, seven symptoms had means greater than 4.0 and are included in the table. Of these symptoms, the highest rated was "complains of being misunderstood or

unappreciated by others," a criterion for Passive-Aggressive (Negativistic) Personality Disorder.

Reliability. Interrater agreement of the symptom ratings was calculated using intraclass correlation (ICC) with the raters treated as variables and ratings for each of the symptoms as cases. Overall, results suggested a moderate level of interrater reliability for the symptom ratings (ICC = .44).

Severity, Prognosis, and Treatment Response

On average, participants rated the severity of the case as moderately high (M = 5.11, SD = .98 on a scale of 1 = very mild to 7 = very severe). Participants rated the prognosis of the patient depicted in the vignette as moderately poor (M = 5.30, SD = 1.17 on a scale of 1 = very good to 7 = very poor) and were not very optimistic about the likelihood of the patient benefiting from therapy (M = 5.03, SD = 1.31 on a scale of 1 = very likely to 7 = very unlikely).

IMI Ratings

Table J.1 (Appendix J) presents the highest rated IMI items (M > 2.5), and the average rating for the Dominant and Hostile scales. Although the sum of items is commonly used for the individual IMI scales, an average was used due to missing ratings for some IMI items for a number of participants. Seven of the 13 items from the Hostile scale had mean ratings of greater than 2.5 on the four-point scale, with the highest rated item being, "when I'm with this person, it appears to me than she is carrying a grudge." Only three of the 11 items from the Dominant scale received a mean rating greater than 2.5, with the highest being, "when I'm with this person, it appears to me that she wants to be the center of attention." Average ratings for the Hostile scale (M = 2.30; SD = .49)

were significantly higher than for the Dominant scale ($M = 2.16$; $SD = .54$), $t(97) = -3.24$, $p < .001$, although both were in the moderate range. Finally, a significant relationship was found to exist between the two scales of the IMI, $r(96) = .68$, $p < .001$.

Reliability. Interrater agreement of the IMI items was calculated using intraclass correlation (ICC) with the raters treated as variables and IMI items as cases. Overall, results suggested a moderately low level of agreement for items on each of the two scales (Dominant scale: ICC = .31, Hostile scale: ICC = .23), suggesting variability between participants in their NER to the case.

Participant Characteristics

Years of Clinical Experience. Correlational and between-group analyses were used to examine the relationship between years of clinical experience and the key variables. Correlational analyses showed there was not a significant relationship between years of clinical experience and a diagnosis of BPD, $r(96) = -.11$, $p = .27$, or BPD representativeness ratings, $F(2, 97) = .05$, $p = .63$. When participants were divided into three groups based on number of years of experience (i.e., less than 10 years, 10 to 20 years, more than 20 years), there was not a significant difference between the groups in frequency of BPD diagnosis, $\chi^2 (2, N = 97) = .94$, $p = .63$, or BPD representativeness ratings, $F(2, 97) = .28$, $p = .76$. In addition, only one BPD symptom was found to be significantly related to years of clinical experience (see Table K.1, Appendix K). No significant correlations were found between years of clinical experience and the mean score on the Dominant scale, $r(96) = -.01$, $p = .93$, or Hostile scale, $r(96) = .14$, $p = .18$, and only three of the 24 IMI items were shown to have a relationship with years of clinical experience (see Table K.2).

Table 4

Mean Representativeness Ratings Given to Axis I and II Diagnoses for Case One, with Intraclass Correlations (ICC)

Axis I Diagnosis	Mean (SD)	Axis II Diagnosis	Mean (SD)
Intermittent Explosive Disorder	4.17 (1.85)	Borderline PD	4.73 (2.06)
Bipolar II	2.29 (1.57)	Narcissistic PD	3.99 (2.02)
Bipolar I	2.19 (1.57)	Histrionic PD	2.76 (1.67)
Dysthymic Disorder	2.11 (1.42)	PAPD	2.60 (1.73)
Major Depressive Disorder	1.97 (1.11)	Antisocial PD	2.41 (1.63)
Adjustment Disorder	1.70 (1.13)	Paranoid PD	2.27 (1.35)
		Dependent PD	1.54 (.93)
		OCPD	1.53 (1.07)
		Depressive PD	1.35 (.87)
		Schizoid PD	1.24 (.59)
		Schizotypal PD	1.24 (.59)
		Avoidant PD	1.22 (.68)
ICC	.46		.50

Note. All ratings were based on a scale from 1 (not at all representative) to 7 (very representative). PAPD = Passive-Aggressive (Negativistic) Personality Disorder. OCPD = Obsessive-Compulsive Personality Disorder.

Finally, no significant relationship was found between years of clinical experience and severity or prognosis ratings, or ratings of the degree to which the person would be likely to benefit from treatment.

Percent of direct patient contact. Correlational and between-group analyses were also used to examine the relationship between the percent of time participants reported spending in direct patient contact and the key variables. It is important to note that approximately one-fourth of the sample did not provide data for this variable, so the following analyses may have been affected by the small sample size. First, there was not a significant relationship between the percent of direct patient contact and a diagnosis of BPD, $r(69) = .06$, $p = .60$, or BPD representativeness ratings, $r(69) = .19$, $p = .14$. When participants were divided into four groups based on percent of direct patient contact (i.e., 0 – 24%, 25 – 49%, 50 – 74%, and 75 – 100%), no significant differences were found for a diagnosis of BPD, χ^2 (3, $N = 71$) = 1.10, $p = .77$, or BPD representativeness ratings, F (3, 67) = 1.09, $p = .36$. Three of the nine BPD symptoms were found to have a relationship with percent of direct patient contact (see Table K.1). No significant correlations were found between percent of direct contact and the mean for the Dominant scale, $r(69) = -.04$, $p = .76$, or the Hostile scale, $r(69) = .20$, $p = .10$, and only five IMI items were found to have a relationship with percent of direct contact (see Table K.2). Finally, no significant relationship was found between percent of direct patient contact and ratings of severity or prognosis, or ratings of the degree to which the person would be likely to benefit from treatment.

Gender. Although hypotheses were not formulated, gender differences were also examined due to the previous finding of clinician gender affecting the diagnosis of BPD

(i.e., Morey & Ochoa, 1989). A chi-square analysis showed female participants were significantly more likely to assign a diagnosis of BPD to the case than male participants (67.31% *vs.* 39.13%), χ^2 (1, N = 98) = 7.80, p = .005. However, a one-way ANOVA showed that there were no significant differences between male and female participants in BPD representativeness ratings (M = 4.52, SD = 2.13 *vs.* M = 4.92, SD = 2.00), and no gender differences were found for BPD symptom ratings. Further, no significant differences were observed between male and female participants for the two IMI scales. Finally, no significant differences between male and female participants were found with respect to ratings of severity or prognosis, or ratings of the degree to which the person would be likely to benefit from treatment.

Order Effects

Order of Diagnostic and Symptom Ratings. The effect of the order in which the diagnostic and symptom ratings were presented was examined because it was hypothesized that those who were asked to assign diagnoses and then rate symptoms (i.e., prototype approach) would be more likely to over-diagnose BPD than those asked to rate symptoms before assigning the diagnosis (i.e., DSM-IV approach). In contrast to the hypothesis, participants who received the symptom ratings first assigned more diagnoses of BPD (60.37%) than those who received the diagnostic ratings first (40.0%); however, the difference failed to reach statistical significance, χ^2 (1, N = 98) = 3.21, p = .073. Similarly, participants who received the symptom ratings first rated BPD as being somewhat more representative of the case (M = 5.08, SD = 1.86) than those who received the diagnostic ratings first (M = 4.36, SD = 2.21), but an ANOVA revealed that the difference also was not significant, F (1, 97) = 3.03, p = .085. Diagnostic and symptom

rating order significantly affected ratings on three of the nine BPD symptoms (see Table L.1 in Appendix L), with one rated higher when symptom ratings were presented first and the other two rated higher when diagnostic ratings were presented first. The order of the diagnostic and symptom ratings did not significantly affect ratings of the IMI Dominant and Hostile scales, ratings of severity or prognosis ratings, or ratings of the degree to which the person in the case would be likely to benefit from treatment.

Order of case. To determine if the order in which the cases were presented to participants influenced ratings for Case One, a series of analyses were conducted comparing those who received Case One first versus second on the key variables. The percentage of participants who selected a diagnosis of BPD for the case did not differ significantly by the order in which they received the case, χ^2 (1, N = 98) = .412, p > .05, (57.44% *vs.* 50.98%). However, a one-way ANOVA showed that participants who received Case One first rated that case as significantly more representative of BPD than participants who received Case One second, F (1, 97) = 9.77, p = .002, (M = 5.38, SD = 1.74 *vs.* M = 4.14, SD = 2.16). There were also order effects on four of the nine BPD symptoms (see Table L.2), with higher ratings when Case One was presented first rather than second. The order of the cases did not significantly impact the mean rating for either of the two IMI scales, participants' ratings of severity or prognosis, or ratings of the degree to which the person depicted in the case would be likely to benefit from treatment.

Relationship Between IMI Scales and Key Variables

Median-split analyses. The ratings for the Dominant and Hostile scales were used to divide participants into two groups (high and low) based on the median scores for each

scale. This technique allowed for between-group analyses for each of the key variables. Results of a chi-square analysis showed that participants' diagnosis of BPD did not differ significantly based on whether they were in the "high" or "low" dominant group, χ^2 (1, N = 98) = 1.52, p = .218, or the "high" or "low" hostile group, χ^2 (1, N = 98) = .241, p = .624. Moreover, a one-way ANOVA showed that BPD representativeness ratings did not differ significantly based on "high" versus "low" Dominant scale scores, F (1, 96) = .015, p = .902, or Hostile scale scores, F (1, 96) = .845, p = .360.

Results were also examined for Narcissistic Personality Disorder (NPD) because it was the second most frequently diagnosed and second highest rated Axis II diagnosis. No significant differences were found in assigned diagnoses of NPD between participants scoring "high" or "low" on the Dominant scale, χ^2 (1, N = 98) = .01, p = .91. However, participants scoring "high" on the Hostile scale were significantly *less* likely to assign a diagnosis of NPD (37.5%) than those scoring "low" on the scale (62.5%), χ^2 (1, N = 98) = 3.98, p = .05. In contrast, participants scoring "high" on the Dominant scale assigned significantly higher NPD ratings (M = 4.90, SD = 1.82) compared to those scoring "low" on the scale (M = 3.04, SD = 1.79), F (1, 96) = 22.99, p < .001. Similarly, NPD ratings for those scoring "high" on the Hostile scale were significantly higher (M = 4.63, SD = 1.96) than for those scoring "low" on the scale (M = 3.20, SD = 1.84), F (1, 96) = 13.59, p < .001.

Regression Analyses. The primary hypotheses were tested using a series of regression analyses examining the degree to which IMI ratings (i.e., Dominant and Hostile scales) predicted a diagnosis of BPD and BPD representativeness ratings. In addition, the analyses were repeated including average BPD symptom ratings as a

predictor since it was hypothesized that clinicians high in NER would base their diagnosis and diagnostic representativeness ratings less on the presence of BPD symptoms in the case than those low in NER (i.e., over-diagnose BPD relative to their symptom ratings). Due to the significant effects of clinician gender on BPD diagnosis and case order on BPD representativeness ratings, the analyses were also conducted using these variables as predictors. Although Hypothesis Two concerned the effects of clinician variables (i.e., years of experience, percent of time in direct patient contact) on BPD ratings and diagnosis, these variables were not significantly related to a BPD diagnosis or BPD representativeness ratings (as noted above) and therefore were not used as predictors. Similarly, because the proposed effects of order of rating of symptoms and diagnoses were not found, this variable was not included in the analyses.

Table M.1 in Appendix M presents the correlations between the primary criterion variables and the IMI scales. The Dominant scale was positively related to average BPD symptom ratings (i.e., the mean rating for the nine symptoms of BPD) and NPD representativeness ratings. The Hostile scale was positively related to BPD representativeness ratings, average BPD symptom ratings, and NPD representativeness ratings.

A logistic regression analysis revealed that the predictive value of a model using Dominant and Hostile scale scores as predictors of a BPD diagnosis was not significant, χ^2 (2, N = 98) = 4.51, p = .105. Overall, the model correctly predicted 57.1% of the diagnoses of BPD (37.8% of the cases not diagnosed as BPD and 73.6% of the cases diagnosed as BPD). This finding represents only a 3% increase in prediction over classification without using Dominant and Hostile scale scores (i.e., base rates). Table

M.2 shows the coefficients, associated Wald statistics, and significance levels for each of the predictors in the model. The Dominant scale was a significant predictor of BPD diagnosis; however, contrary to expectation, higher Dominant scale scores predicted *fewer* BPD diagnoses. This result should be viewed with caution, however, since the Dominant scale was not significantly correlated with BPD diagnosis (see Table M.1) nor was it a significant predictor of BPD diagnosis when considered individually (as seen in the first step in the regression).

When the regression was repeated controlling for average BPD symptom ratings, the predictive value of the model remained non-significant, χ^2 (2, N = 98) = 5.66, p = .129. Overall, this model correctly predicted 60.2% of the diagnoses (44.4% of the cases not diagnosed as BPD and 73.6% of the cases diagnosed as BPD). This represents only a 6% increase in prediction over a model that does not include the Dominant and Hostile scales or the average BPD symptom ratings (i.e., base rates). Table M.3 shows the coefficients, associated Wald statistics, and significance levels for each of the predictors in the model. The Dominant scale remained significant when controlling for average BPD symptom ratings; however, the relationship was in the opposite direction than expected and should be viewed with caution (as noted above). For both models, the relationship between Hostile scale scores and BPD diagnosis was non-significant.

A final logistic regression was conducted adding clinician gender as a predictor of BPD diagnosis due to the significant difference in the frequency that men and women diagnosed BPD. Table M.4 shows the coefficients, associated Wald statistics, and significance levels for each of the predictors in the model. Although Dominant scores

remained significant, gender emerged as a highly significant predictor of BPD diagnosis, with women being more likely to assign the diagnosis.

With regard to BPD representativeness ratings, an initial multiple regression analysis using Dominant and Hostile scale scores as predictors revealed that the overall model had a significant predictive value, R = .28 (R^2 = .08, Adjusted R^2 = .06); F (2,97) = 3.95, p = .02. Table N.1 in Appendix N shows the coefficients, t-values, and significance levels for the individual predictors in the model. Hostile scale scores predicted higher BPD representativeness ratings with scores on the Hostile scale uniquely accounting for 7.51% of the variance in BPD representativeness ratings.

In order to test the degree to which the IMI scales predicted BPD representativeness ratings beyond BPD symptom ratings, a follow-up analysis was conducted that included average BPD symptom ratings as a predictor. Overall, this multiple regression model had a significant predictive value, R = .43 (R^2 = .19, Adjusted R^2 = .16); F (2,97) = 7.21, $p < .001$. Table N.2 shows the coefficients, t-values, and significance levels of the individual predictors in this model. The Hostile scale became only a marginally significant predictor of BPD representativeness ratings when average rating of BPD symptoms was entered as a predictor, thereby reducing the amount of variance explained by Hostile scale ratings to 3.5%. As would be expected, average BPD symptom ratings were a highly significant predictor of BPD representativeness ratings, uniquely accounting for 12.0% of the variance in this model. Overall, this analysis revealed that the Hostile scale ratings did not account for a significant portion of the variance above what was explained by the DSM-IV symptom ratings.

To determine the impact of case order, the order in which the participants received the cases was entered as a predictor in the model. Overall, this multiple regression model had a significant predictive value, R = .47 (R^2 = .22, Adjusted R^2 = .19); F (2,97) = 6.54, $p < .001$. Table N.3 shows the coefficients, t-values, and significance levels of the individual predictors in this model. When the effect of case order was taken into account, the predictive value of the Hostile scale declined even further, thereby reducing the amount of variance explained by the Hostile scale ratings to 2.0%. Additionally, the predictive value of the BPD symptoms also declined, reducing the amount of variance explained by this variable from 12.0% to 9.5%, although it was still the most significant predictor of BPD representativeness ratings. Finally, case order was a significant predictor, with higher BPD representativeness ratings when Case One was presented first.

Given that accounting for average BPD symptom ratings substantially reduced the amount of variance explained by the IMI scales in the above models, a multiple regression analysis was conducted that used the Dominant and Hostile scales to predict BPD symptom ratings. Overall, this model had a significant predictive value, R = .35 (R^2 = .13, Adjusted R^2 = .11); F (2,97) = 6.82, $p < .01$. Table N.4 shows the coefficients, t-values, and significance levels of the individual predictors in this model. Hostile scale ratings were found to predict average BPD symptom ratings, with increasing Hostile scale ratings associated with increasing BPD symptom ratings. Dominant scale scores were found to be non-significant. Overall, these findings suggest that BPD symptom ratings mediate the relationship between Hostile scale scores and BPD diagnostic ratings. In other words, it appears that ratings of increased hostility towards the case are

predictive of increased BPD symptom ratings which in turn predict increased BPD diagnostic ratings.

Finally, regression analyses were conducted examining the degree to which Dominant and Hostile scale scores predict a diagnosis of Narcissistic Personality Disorder (NPD) and NPD representativeness ratings. As noted above, both of the IMI scales were significantly correlated with NPD representativeness ratings but not NPD diagnoses (see Table N.1). A logistic regression analysis using Dominant and Hostile scale scores as predictors of diagnosis of NPD, was not significant, χ^2 (2, N = 98) = 1.46, p = .48, and neither the Dominant nor Hostile scales were shown to be significant predictors of NPD diagnosis. Overall, the model correctly predicted 100% of the cases not diagnosed with NPD but failed to predict any of the cases diagnosed with NPD. This finding does not represent a significant increase in prediction over classification without using the Dominant and Hostile Scales (i.e., base rates). Table O.1 in Appendix O shows the coefficients, associated Wald statistics, and significance levels for each of the predictors in the model. Next, a multiple regression analysis was conducted to test the predictive value of the Dominant and Hostile scales for representativeness ratings of NPD. Overall, this model had a significant predictive value, R = .44 (R^2 = .20, Adjusted R^2 = .18); F (2,97) = 11.69, $p < .001$. Table O.2 shows the coefficients, t-values, and significance levels of the individual predictors in this model. Ratings on the Dominant scale were shown to be highly predictive of NPD representativeness ratings, and uniquely accounted for 8.0% of the variance in the ratings. Ratings on the Hostile scale were shown to be a non-significant predictor of NPD ratings.

Relationship of IMI Scales with ratings of Severity, Prognosis, and Treatment Benefit

Table P.1 in Appendix P shows the relationship between the two IMI scales and ratings of severity, prognosis, and the likelihood the person depicted in the case would benefit from treatment. As predicted, Dominant scale scores were correlated with ratings of severity for the case, and Hostile scale scores were correlated with ratings of severity, prognosis, and the likelihood that the person depicted in the case would benefit from treatment. Note that higher scores indicated worse prognosis and treatment response.

Table P.2 shows the results of a series of multiple regression analyses examining the degree to which the IMI scales predict ratings of severity, prognosis, and the likelihood the person depicted in the case would be to benefit from treatment. Contrary to the hypothesis, ratings on the Dominant and Hostile scales were not significant predictors of severity ratings. Hostile scale scores were shown to be a significant predictor of prognosis and likelihood of treatment benefit, with higher Hostility predicting ratings of worse prognosis and lower likelihood of the patient benefiting from treatment.

Over- and Under-diagnostic Bias

The final analyses examined within-subject diagnostic bias with respect to BPD by comparing participants' "criterion diagnosis" based on their symptom ratings with their assigned diagnosis. Those who endorsed an insufficient number of BPD symptoms but still diagnosed BPD were considered to have engaged in "over-diagnostic bias" (N = 36). Those who endorsed a sufficient number of BPD symptoms but failed to assign the diagnosis were said to have engaged in "under-diagnostic bias" (N = 12). Overall, 48

participants (50% of the sample) demonstrated inconsistency between their symptom ratings and their assigned diagnosis.

It had been hypothesized that NER would have a significant effect on diagnostic bias, specifically an over-diagnosis of BPD. A logistic regression analysis using Dominant and Hostile scale scores as predictors of over-diagnostic bias showed that the predictive value of the overall model was not significant, χ^2 (2, N = 98) = 3.56, p = .17. Neither IMI scale was a significant predictor of over-diagnostic bias. Overall, correct classification of the model was 66.3%, with the model correctly classifying 96.8% of the cases not showing over-diagnostic bias but only 13.9% of the cases that showed over-diagnostic bias. This finding represents only a 3% increase in prediction over classification without using Dominant and Hostile scale scores (i.e., base rates). Table Q.1 in Appendix Q shows the coefficients, associated Wald statistics, and significance levels for each of the predictors in the model.

With regard to under-diagnostic bias, a logistic regression analysis using Dominant and Hostile scale scores as predictors of under-diagnostic bias showed that the predictive value of the overall model was not significant, χ^2 (2, N = 98) = 4.27, p = .12. Again, neither IMI scale was a significant predictor. Overall, the model correctly predicted 87.8% of the cases of under-diagnostic bias, correctly classifying 100% of the cases showing no under-diagnostic bias and 0% of the cases showing over-diagnostic bias. This finding does not represent an increase in prediction over classification without using Dominant and Hostile scale scores (i.e., base rates). Table Q.2 shows the coefficients, associated Wald statistics, and significance levels for each of the predictors in the model.

Between-group analyses were conducted to examine the effect of order of diagnostic and symptom ratings on over- and under-diagnostic bias since it was hypothesized that subjects who rated symptoms first (DSM-IV approach) would be more reliant on the DSM criteria, whereas those who assigned diagnoses first would utilize a prototype approach and would be more influenced by NER. Contrary to this hypothesis, participants who received the diagnostic ratings first were significantly *less* likely to engage in over-diagnostic bias (25.53%) than those who received the symptom ratings first (47.06%), χ^2 (1, N = 98) = 4.88, p = .03. With regard to under-diagnostic bias, participants who received the diagnostic ratings first were significantly more likely to engage in under-diagnostic bias (19.15%) than participants who received the symptom ratings first (5.88%), χ^2 (1, N = 98) = 4.01, p = .05. These results were supported by correlational analyses; there was a significant relationship between order of symptom and diagnostic ratings and over-diagnosis, r(96) = .22, p = .03, and under-diagnosis, r(96) = -.20, p < .05. In addition, logistic regressions using the IMI scales and order of symptom and diagnostic ratings as predictors revealed that only rating order was a significant predictor of over-diagnostic and under-diagnostic biases (see Table Q.3). Both models were significant, χ^2 (3, N = 98) = 8.95, p = .03, and χ^2 (3, N = 98) = 8.95, p = .03, respectively. However, they represented only a 4% and 2% increase in accuracy of classification.

Finally, given the finding that female clinicians assigned BPD more frequently for Case One and the previous finding (Morey & Ochoa, 1989) that female clinicians were more likely to over-diagnose BPD, logistic regressions were conducted using participant gender and the IMI scales as predictors of over- and under-diagnostic bias. The results of

these analyses showed that only clinician gender was a significant predictor of over-diagnostic bias and under-diagnostic bias, with female clinicians more likely to over-diagnose BPD and male clinicians more likely to under-diagnose the disorder (see Table Q.4). Although both models were significant, χ^2 (3, N = 98) = 8.01, p = .05, and χ^2 (3, N = 98) = 8.68, p = .03, they resulted in only a 2% and 1% increase in correct classification over base rates, respectively.

Case Two

Diagnosis and Confidence Ratings

Table 5 presents the rates at which diagnoses were assigned for Axis I and II. For Axis I, Dysthymic Disorder was the most frequently assigned diagnosis, with 44.9% of the participants assigning it as the Axis I diagnosis that was most representative of the case. For Axis II, Dependent PD was the most frequent diagnosis assigned, with 43.9% of the respondents assigning it as the Axis II diagnosis most representative of the case. BPD was the next most frequently assigned diagnosis for Axis II, with 34.7% of participants assigning it as the most representative diagnosis. Overall, participants reported having a high level of confidence in their Axis II diagnosis (M =5.06, SD = 1.55), but a low to moderate level of confidence in their Axis I diagnosis (M = 2.97, SD = 1.58).

Reliability. Inter-rater agreement of the Axis I and II diagnoses assigned to the case was examined using the *kappa* statistic (Cohen, 1960). *Kappa* values for Axis I and Axis II diagnoses indicated a low to moderate level of agreement (see Table 5).

Diagnostic Representativeness Ratings

Participants also rated the representativeness of a series of Axis I and II diagnoses for the case. Table 6 lists the average representativeness rating given to each diagnosis by the participants. For Axis I, Dysthymic Disorder received the highest representativeness rating, in the low to moderate range, with other Axis I diagnoses rated low in representativeness for the case. Dependent PD received the highest average representativeness rating on Axis II, rated moderate to high in representativeness for the case, followed by BPD, which was also in the moderate to high range.

Reliability. Interrater agreement of the representativeness ratings was calculated using intraclass correlation (ICC). The raters were treated as variables and ratings as cases. The ICC provides an estimate of the interrater reliability based on the proportion of variance associated with differences between the measured ratings and variability among the raters (Nichols, 1998). Overall, results suggested a low level of interrater reliability for the representativeness ratings for Axis I diagnoses and a moderate level of reliability for the Axis II diagnoses (see Table 6).

Symptom Ratings

Table I.2 (Appendix I) lists the average symptom ratings for all nine symptoms of BPD as well as additional highly rated symptoms (i.e., mean rating of 4.0 or higher). Three of the nine symptoms of BPD received a mean greater than 4.0, with the most highly rated symptoms being the two intended symptoms for the case: "frantic efforts to avoid real or imagined abandonment" and "affective instability due to a marked reactivity of mood." The remaining symptoms of BPD, which were not intended to be present in

Table 5

Axis I and II Diagnoses Assigned to Case Two, with Inter-rater Agreement (Kappa)

Axis I Diagnosis	Freq (%)	Axis II Diagnosis	Freq (%)
Dysthymic Disorder	44 (44.9)	Dependent PD	43 (43.9)
Adjustment Disorder	18 (18.4)	Borderline PD	34 (34.7)
Bipolar II Disorder	14 (14.3)	Narcissistic PD	10 (10.2)
Major Depressive Disorder	7 (7.1)	Depressive PD	5 (5.1)
Intermittent Explosive Disorder	7 (7.1)	Histrionic PD	4 (4.1)
Bipolar I Disorder	5 (5.1)	OCPD[a]	1 (1.0)
No Diagnosis	3 (3.1)	PAPD	1 (1.0)
Kappa	.34		.39

Note. "No Diagnosis" includes participants who either skipped the question in error or did not want to assign a diagnosis to the case. Diagnoses that were not assigned by any participants are not included in the table. OCPD = Obsessive-Compulsive Personality Disorder. PAPD = Passive-Aggressive (Negativistic) Personality Disorder.

the case, had means of less than 4.0. In addition to the BPD symptoms, seven symptoms had means greater than 4.0 and are included in the table. Of these symptoms, the highest rated was: "marked distress which is in excess of what would be expected from the stressor," a criterion for Adjustment Disorder.

Reliability. Interrater agreement of the symptom ratings was calculated using intraclass correlation (ICC) with the raters treated as variables and ratings for each of the

symptoms as cases. Overall, results suggested a moderate level of interrater reliability for the symptom ratings (ICC = .44).

Severity, Prognosis, and Treatment Response

On average, participants rated the severity of the case as moderately high (M = 4.67, SD = 1.08 on a scale of 1 = very mild to 7 = very severe). They also rated the prognosis of the patient depicted in the vignette as moderately poor (M = 5.00, SD = 1.27 on a scale of 1 = very good to 7 = very poor), and rated the likelihood of the patient benefiting from therapy similarly (M = 4.60, SD = 1.44 on a 7-point scale of 1 = very likely to 7 = not very likely).

IMI Ratings

Table J.2 (in Appendix J) presents the highest rated IMI items ($M > 2.5$), and the average rating for the Dominant and Hostile scales. Like Case One, an average was used due to missing ratings for a number of participants. Two of the 13 items from the Hostile scale contained means of greater than 2.5 on the four-point scale, with the highest rated item being "when I'm with this person, she makes me feel forced to shoulder all the responsibility." Only one of the 11 items from the Dominant scale received a mean rating greater than 2.5: "when I'm with this person, it appears to me that she wants to be the center of attention." Average scale ratings were somewhat higher for the Hostile scale (M = 1.95; SD = .48) compared to the Dominant scale (M = 1.87; SD = .59), but not significantly different, $t(97) = -1.68$, $p = .09$, and both were in the moderate to low range. Finally, a significant relationship was found between the two scales, $r(96) = .63$, $p < .001$.

Table 6

Mean Representativeness Ratings Given to Axis I and II Diagnoses for Case Two, with Intraclass Correlations (ICC)

Axis I Diagnosis	Mean (SD)	Axis II Diagnosis	Mean (SD)
Dysthymic Disorder	2.63 (1.69)	Dependent PD	4.86 (1.71)
Major Depressive Disorder	2.21 (1.41)	Borderline PD	4.46 (2.06)
Adjustment Disorder	2.18 (1.52)	Histrionic PD	3.55 (1.75)
Bipolar II Disorder	1.87 (1.47)	Narcissistic PD	3.25 (1.96)
Bipolar I Disorder	1.69 (1.34)	PAPD	2.62 (1.64)
IED	.59 (1.11)	Depressive PD	2.30 (1.48)
		Paranoid PD	1.77 (1.10)
		Avoidant PD	1.57 (1.14)
		Antisocial PD	1.39 (.82)
		OCPD	1.35 (.79)
		Schizoid PD	1.24 (.73)
		Avoidant PD	1.20 (.56)
ICC	.06		.55

Note. All ratings were based on a scale from 1 (not at all representative) to 7 (very representative). PAPD = Passive-Aggressive (Negativistic) Personality Disorder. OCPD = Obsessive-Compulsive Personality Disorder. IED = Intermittent Explosive Disorder.

Reliability. Interrater agreement of the IMI items was calculated using intraclass correlation (ICC) with the raters treated as variables and items as cases. Overall, results

suggested a low to moderate level of agreement for both the Dominant (ICC = .31) and the Hostile (ICC = .36) scales, suggesting variability in the way participants responded to the case.

Participant Characteristics

Years of Clinical Experience. Correlational and between-group analyses were used to examine the relationship between years of clinical experience and the key variables. The results of correlational analyses showed there was no significant relationship between years of clinical experience and a diagnosis of BPD, $r(96) = -.02$, $p = .85$, or BPD representativeness ratings, $r(96) = .05$, $p = .65$. When participants were divided into three groups based on years of experience (i.e., less than 10 years, 10 to 20 years, more than 20 years), no significant difference was found between the groups in the frequency of diagnosing BPD, χ^2 (2, $N = 97$) = 4.73, $p = .09$, or BPD representativeness ratings, F (2, 97) = 1.10, $p = .34$. Only two BPD symptoms were found to be related to years of experience (see Table K.3). No significant correlations were found between years of clinical experience and the Dominant scale, $r(96) = .01$, $p = .91$, the Hostile scale, $r(96) = .14$, $p = .18$, or any of the 24 IMI items. A significant relationship was found between years of clinical experience and ratings of prognosis for the case, with ratings of poorer prognosis with increasing years of experience, $r(96) = .24$, $p = .02$. However, no significant relationship was found between years of clinical experience and ratings of severity, or ratings of the degree to which the person depicted in the case would be likely to benefit from treatment.

Percent of direct patient contact. Correlational and between group analyses were also used to examine the relationship between the percent of time participants' reported

spending in direct patient contact and the key variables. As noted previously, approximately one fourth of the sample did not provide data for this variable, so the following analyses should be viewed with some caution. There was not a significant relationship between the percent of direct patient contact and a diagnosis of BPD, $r(69) = .01$, $p = .95$, or BPD representativeness ratings, $r(69) = -.07$, $p = .54$. When participants were divided into four groups based on their percent of time in direct patient contact (i.e., 0 – 24%, 25 – 49%, 50 – 74%, and 75 – 100%), no significant differences were found for a diagnosis of BPD, χ^2 (3, $N = 71$) = .85, $p = .84$, or BPD representativeness ratings, F (3, 68) = .137, $p = .94$. With respect to the BPD symptom ratings, two of the nine symptoms were found to have a relationship with percent of direct patient contact (see Table K.3). Although no significant relationship was found between percent of direct contact and the Dominant scale, $r(69) = 04$, $p = .73$, ratings on the Hostile scale were found to increase in relation to increasing time spent in direct patient contact, $r(69) = .27$, $p = .02$. Of the individual IMI items, three were found to have a relationship with percent of direct contact (see Table K.4). Finally, a significant relationship was found between percent of direct contact and expected treatment benefit, with lower ratings of likelihood of benefiting from therapy associated with a higher percent of time spent in direct patient contact, $r(69) = .25$, $p = .03$. No significant relationship was found between percent of direct patient contact and ratings of severity or prognosis.

Gender. Although more women than men assigned a diagnosis of BPD (42.31% *vs.* 26.09%), the difference was not statistically significant, χ^2 (1, N = 98) = 2.83, $p = .09$. Similarly, a one-way ANOVA showed that there were no significant differences between male and female participants in BPD representativeness ratings ($M = 4.51$, $SD = 2.00$ and

M = 4.42, SD = 2.14, respectively). With regard to the symptom ratings, no significant differences were observed between male and female participants, although ratings of "frantic efforts to avoid real or imagined abandonment" was marginally significant, with women rating that symptom somewhat higher than men, F (1, 97) =3.77, p = .06 (M = 6.15, SD = 1.14 vs. M = 5.70, SD = 1.19). No significant differences were observed between male and female participants for the two IMI scales. Finally, no significant differences between male and female participants were found for ratings of severity or prognosis, or ratings of the degree to which the person depicted in the case would be likely to benefit from treatment.

Order Effects

Order of Diagnostic and Symptom Ratings. The effect of the order in which the diagnostic and symptom ratings were presented was also examined because it has been hypothesized that those asked to assign diagnoses before symptom ratings (i.e., prototype approach) would be more likely to diagnose BPD than those asked to rate symptoms first (i.e., DSM-IV approach). Results failed to support the hypothesis, with no significant difference in diagnosis of BPD or BPD representativeness ratings between participants who received the symptom ratings first versus those who received the diagnostic ratings first. Order also did not affect any of the BPD symptom ratings. Finally, the order of the diagnostic and symptom ratings did not significantly affect ratings of the IMI Dominant and Hostile scales, ratings of severity or prognosis, or ratings of the degree to which the person would be likely to benefit from treatment.

Order of case. To determine if the order in which the cases were presented influenced ratings for Case Two, a series of analyses were conducted comparing those

who received Case Two first versus second on the key variables. Participants who received Case Two first were significantly more likely to assign a diagnosis of BPD (52.94%) than those who received Case Two second (14.89%), χ^2 (1, N = 98) = 15.63, p < .001. Similarly, a one-way ANOVA showed that participants who received Case Two first rated that case as significantly more representative of BPD than participants who received Case Two second, F (1, 97) = 12.06, p = .001, (M = 5.12, SD = 1.81 *vs.* M = 3.74, SD = 2.10). There were also order effects of case on three of the nine BPD symptoms (see Table L.3), with higher ratings when Case Two was presented first rather than second. A significant difference was observed for one of the two BPD symptoms intended to be present in the case (i.e., affective instability).

With regard to the IMI ratings, the order of the cases significantly impacted the mean for the Dominant scale, with participants assigning higher ratings when Case Two was presented first compared to when it was presented second, F (1, 97) = 17.21, p < .001 (M = 2.09, SD = .61 *vs.* M = 1.63, SD = .46). Similarly, the order of the cases impacted the mean for the Hostile scale, with higher ratings when Case Two was presented first compared to when it was presented second, F (1, 97) = 10.17, p = .002 (M = 2.10, SD = .50 *vs.* M = 1.80, SD = .41). Additionally, the order of the cases significantly impacted participants' ratings of confidence in their Axis II diagnosis, with participants reporting higher levels of confidence when Case Two was presented first compared to when it was presented second, F (1, 97) = 8.78, p = .004 (M = 5.49, SD = 1.40 *vs.* M = 4.60, SD = 1.60). However, case order did not have a significant effect of ratings of confidence in Axis I diagnoses. Case order also significantly impacted ratings of prognosis, with participants who received Case Two first assigning lower prognosis

ratings compared to those who received the case second, F (1, 97) = 13.95, $p < .001$ (M = 5.43, SD = 1.03 *vs.* M = 4.53, SD = 1.35). However, case order did not significantly impact participants' ratings of severity or the degree to which the person depicted in the case would be likely to benefit from treatment.

Relationship Between IMI Scales and Key Variables

Median-split analyses. As in Case One, the ratings for the Dominant and Hostile scales were used to divide participants into two groups (high and low) based on the median scores for each scale to allow for between-group analyses for each of the key variables. A chi-square analysis showed that participants who rated the case as "high" on the Dominant scale were significantly more likely to assign a diagnosis of BPD (50.00%) compared to those who were "low" on the scale (17.39%), χ^2 (1, N = 98) = 11.46, p = .001. However, diagnoses of BPD did not differ significantly for those in "high" and "low" Hostile scale groups, χ^2 (1, N = 98) = 1.94, p = .16. A one-way ANOVA revealed that BPD representativeness ratings were also significantly higher for participants in the "high" Dominant scale group compared to the "low" group, F (1, 96) = 24.43, $p < .001$ (M = 5.33, SD = 1.68 *vs.* M = 3.47, SD = 2.03). A similar pattern was found for the Hostile scale, with participants scoring "high" on the scale assigning significantly higher BPD representativeness ratings (M = 4.96, SD = 1.84) compared to those scoring "low" on the scale (M = 3.86, SD = 2.17), F (1, 96) = 7.27, p = .01.

The effect of Dominant and Hostile scale ratings was also examined for Dependent Personality Disorder (DPD) since it was the most frequently assigned diagnosis and received the highest representativeness ratings. Results of chi-square analysis revealed that participants in the "low" Dominant group were more likely to

assign a diagnosis of DPD (50.00%) than those in the "high" Dominant group (17.39%), χ^2 (1, N = 98) = 16.03, $p < .001$. However, there were no significant differences between those high versus low on the Hostility scale, χ^2 (1, N = 98) = 2.29, $p = .13$. Moreover, the results of ANOVA showed that there was not a significant difference in DPD representativeness ratings between "high" and "low" groups on both IMI scales.

Regression Analyses. As in Case One, the primary hypotheses were tested using a series of regression analyses examining the degree to which IMI ratings (i.e., Dominant and Hostile scales) predicted a diagnosis of BPD and BPD representativeness ratings. In addition, the analyses were repeated including average BPD symptom ratings as a predictor since it was hypothesized that clinicians high in NER would base their diagnosis and diagnostic representativeness ratings less on the BPD symptoms in the case (i.e., over-diagnosis of BPD relative to their symptom ratings). Due to the significant findings of case order on BPD diagnoses and representativeness ratings, the analyses were also conducted using case order as a predictor. Although Hypothesis Two concerned the effects of participants' years of experience and percent of time spent in direct patient contact, these variables were not significantly related to a BPD diagnosis or BPD representativeness ratings (as noted above) and therefore were not used as predictors. Similarly, order of diagnostic and symptom ratings did not significantly affect BPD diagnosis or representativeness ratings and was not included as a predictor in the regressions. Finally, analyses were conducted for a DPD diagnosis and DPD representativeness ratings because DPD was the most frequently assigned diagnosis and had the highest representativeness ratings.

Table R.1 in Appendix R presents the correlations between the predictors and criterion variables. Both IMI scales were significantly related to a diagnosis of BPD, BPD representativeness ratings, and average BPD symptom ratings. However, the Dominant scale also had a significant inverse relationship to a diagnosis of Dependent Personality Disorder.

A logistic regression analysis revealed that the predictive value of a model using Dominant and Hostile scale scores as predictors of BPD diagnosis was significant, χ^2 (2, N = 98) = 14.61, p = .001. Overall, the model correctly predicted 68.4% of the diagnoses of BPD (85.9% of the cases not diagnosed as BPD and 35.3% of the cases diagnosed as BPD). This finding represents a 3.1% increase in prediction over classification without using Dominant and Hostile scale scores (i.e., base rates). Table R.2 shows the coefficients, associated Wald statistics, and significance levels for each of the predictors in the model. Only the Dominant scale was a significant predictor of BPD diagnosis, with higher ratings on the scale leading to more frequent diagnoses of BPD.

When the regression was repeated controlling for average BPD symptom ratings, the predictive value of the model remained significant, χ^2 (2, N = 98) = 23.83, $p < .001$. Overall, this model correctly predicted 71.4% of the diagnoses (84.4% of the cases not diagnosed as BPD and 47.1% of the cases diagnosed as BPD). This represents a 6.1% increase in prediction over a model that does not include the Dominant and Hostile scales or the average BPD symptom ratings. Table R.3 shows the coefficients, associated Wald statistics, and significance levels for each of the predictors in the model. Although average BPD symptom ratings was a strong predictor of a diagnosis of BPD, the Dominant scale remained a significant predictor when controlling for average BPD

symptom ratings. As with the first model, the relationship between Hostile scale scores and BPD diagnosis was non-significant.

Since participants who received Case Two first were more likely to assign a diagnosis of BPD, a logistic regression was conducted using case order and the IMI scales as predictors of BPD diagnosis. Overall, the predictive value of the model was significant, χ^2 (2, N = 98) = 22.94, $p < .001$, and the model correctly predicted 73.5% of the diagnoses (82.8% of the cases not diagnosed as BPD and 55.9% of the cases diagnosed as BPD). This represents an 8.2% increase in prediction over a model that does not include the Dominant and Hostile scales or the case order. Table R.4 shows the coefficients, associated Wald statistics, and significance levels for each of the predictors in the model. The results of this analysis showed case order to be a highly significant predictor of BPD diagnosis, however, scores on the Dominant subscale continued to be a significant predictor in this model.

With regard to BPD representativeness ratings, an initial multiple regression analysis using Dominant and Hostile scale scores as predictors revealed that the overall model had a significant predictive value, R = .46 (R^2 = .21, Adjusted R^2 = .19); F (2,97) = 12.52, $p < .001$. Table S.1 in Appendix S shows the coefficients, t-values, and significance levels for the individual predictors in the model. Only Dominant scale scores predicted BPD representativeness ratings, with scores on the Dominant scale uniquely accounting for 10.3% of the variance in BPD representativeness ratings.

In order to test the degree to which the IMI scales predicted BPD representativeness ratings beyond BPD symptom ratings, a follow-up analysis was conducted that included average BPD symptom ratings as a predictor. Overall, this

multiple regression model had a significant predictive value, R = .60 (R^2 = .37, Adjusted R^2 = .34); F (2,97) = 17.80, $p < .001$. Table S.2 shows the coefficients, t-values, and significance levels of the individual predictors in this model. Average BPD symptom rating was a highly significant predictor of BPD representativeness ratings, uniquely accounting for 20.0% of the variance in this model. However, the Dominant scale continued to be a significant predictor of BPD representativeness, although the amount of variance uniquely explained by Dominant scale ratings was reduced from 10.3% to 7.5%. Overall, this analysis revealed that the Dominant scale ratings accounted for a significant portion of the variance above what was explained by the DSM-IV symptom ratings.

To determine the impact of case order, the order in which the participants received the cases was entered as another predictor in the model. Overall, this multiple regression model had a significant predictive value, R = .63 (R^2 = .40, Adjusted R^2 = .37); F (2,97) = 15.02, $p < .001$. Table S.3 shows the coefficients, t-values, and significance levels of the individual predictors in this model. Case order was a significant predictor of BPD representativeness ratings, with higher BPD ratings when Case Two was presented first, and BPD symptom ratings remained a strong predictor. The Dominant scale also remained a significant predictor of BPD representativeness ratings, although it uniquely accounted for only 4.8% of the total variance.

Given that accounting for average BPD symptom ratings reduced the amount of variance explained by the IMI scales in the above models, a multiple regression analysis was conducted that used the Dominant and Hostile scales to predict BPD symptom ratings. Overall, this model had a significant predictive value, R = .48 (R^2 = .23, Adjusted R^2 = .21); F (2,97) = 14.05, $p < .01$. Table S.4 shows the coefficients, t-values,

and significance levels of the individual predictors in this model. Higher Hostile scale ratings were found to predict higher BPD symptom ratings, whereas Dominant scale scores were only marginally significant. Overall, these findings suggest that BPD symptom ratings mediate the relationship between Hostile scale scores and BPD diagnostic ratings. In other words, it appears that ratings of increased hostility towards the case predicted increased BPD symptom ratings which in turn predict increased BPD diagnostic ratings. In contrast, Dominant scale scores had a direct effect on BPD representativeness ratings but only a marginally significant effect on BPD symptom ratings.

Finally, regression analyses were conducted examining the degree to which Dominant and Hostile scale scores predict a diagnosis of Dependent Personality Disorder (DPD), and DPD representativeness ratings. As noted above, there was a significant relationship between the Dominant scale and DPD diagnoses but not DPD representativeness ratings and the Hostile scale was not significantly related to either (see Table R.1). A logistic regression analysis revealed that the predictive value of the overall model using Dominant and Hostile scale scores as predictors of DPD diagnosis was highly significant, χ^2 (2, N = 98) = 26.63, $p < .001$. Overall, the model correctly predicted 70.4% of the cases (74.5% of those not diagnosed with DPD and 65% of those diagnosed with DPD). This finding represents a 14.3% increase in prediction over classification without using the Dominant and Hostile Scales (i.e., base rates). Table T.1 in Appendix T shows the coefficients, associated Wald statistics, and significance levels for each of the predictors in the model. Scores on the Dominant scale were highly

predictive of a diagnosis of DPD, with lower scores on the scale predicting a diagnosis of DPD, whereas the Hostile scale was not significant.

Next, a multiple regression analysis was conducted to test the predictive value of the Dominant and Hostile scales for representativeness ratings of DPD. Overall, this model was not significant, R = .15 (R^2 = .02, Adjusted R^2 = .001); F (2,97) = 1.04, p =.36, and neither IMI scale was a significant predictor of DPD representativeness ratings (see Table T.2).

Relationship of IMI Scales with Ratings of Severity, Prognosis, and Treatment Benefit

Table U.1 in Appendix U shows the relationship between the two IMI scales and severity, prognosis, and ratings of the likelihood that person depicted in the case would benefit from treatment. As predicted, Dominant and Hostile scale scores were both significantly related to ratings of severity and prognosis for the case, and the Hostile scale was related to ratings of the likelihood that the person depicted in the case would benefit from treatment. Note that higher scores reflected ratings of poorer prognosis and lower likelihood of the patient benefiting from treatment.

Table U.2 shows the results of a series of multiple regression analyses conducted examining the degree to which the IMI scales predict severity, prognosis, and the likelihood the person depicted in the case would be to benefit from treatment. Consistent with the hypotheses, the Dominant scale was a significant predictor of severity ratings and the Hostile scale was shown to be a significant predictor of ratings of the likelihood that the person depicted in the vignette would benefit from treatment. Higher scores on the IMI scales predicted higher severity and lower likelihood of response to treatment.

However, contrary to prediction, neither IMI scale was a significant predictor of prognosis.

Over- and Under-diagnostic Bias

As in Case 1, the consistency between participants' "criterion diagnosis" based on BPD symptom ratings and their assigned diagnosis was determined to examine the effect of NER on within-subject over- and under-diagnostic bias for BPD. Those who did not endorse a sufficient number of BPD symptoms but still assigned a diagnosis of BPD were considered to have engaged in "over-diagnostic bias" (N = 14). Those who endorsed a sufficient number of BPD symptoms but failed to assign the diagnosis were classified as having engaged in "under-diagnostic bias" (N = 17). Overall, about one-third (31.6%) of the participants demonstrated inconsistency between their symptom ratings and their assigned diagnosis.

A series of logistic regression analyses were conducted to examine to what extent the two IMI scales contributed to these forms of bias. A logistic regression analysis using Dominant and Hostile scale scores as predictors of over-diagnostic bias showed that the predictive value of the overall model was nonsignificant, χ^2 (2, N = 98) = 2.34, p = .31, and neither scale was a significant predictor. Overall, the model correctly predicted 85.7% of the cases of over-diagnostic bias, correctly predicting 100% of the cases not showing over-diagnostic bias and 0% of the cases showing over-diagnostic bias. This finding does not represent an increase in prediction over classification without using Dominant and Hostile scale scores as predictors (i.e., base rates). Table V.1 in Appendix V shows the coefficients, associated Wald statistics, and significance levels for each of the predictors in the model.

A logistic regression analysis using Dominant and Hostile scale scores as predictors of under-diagnostic bias showed that the predictive value of the overall model was also nonsignificant, χ^2 (2, N = 98) = 2.46, p = .29, and neither scale significantly predicted under-diagnostic bias. Overall, the model correctly predicted 82.7% of the cases of over-diagnostic bias (100% of the cases showing no under-diagnostic bias, and 0% of the cases showing over-diagnostic bias), which does not represent an increase in prediction over classification without using Dominant and Hostile scale scores. Table V.2 shows the coefficients, associated Wald statistics, and significance levels for each of the predictors in the model.

Given previous findings that the order of rating symptoms versus diagnoses affected diagnostic bias (e.g., Crosby & Sprock, 2004), a between-groups analysis was conducted to examine the effect of order on over- and under-diagnostic bias. Chi-square analyses showed that there were no significant differences in the frequency of over-diagnostic bias, χ^2 (1, N = 98) = .03, p = .87, or under-diagnostic bias, χ^2 (1, N = 98) = .38, p = .54, based on whether participants rated symptoms first and then assigned diagnoses or assigned diagnoses and then rated symptoms. These results were supported by correlational analyses, as there was no relationship between order of symptom and diagnostic ratings and over-diagnosis, r(96) = -.02, p = .87, or under-diagnosis, r(96) = .06, p = .53. In addition, logistic regression analyses showed that neither of the IMI scales nor the order of diagnostic and symptom ratings were predictive of over- or under-diagnostic bias and neither model was significant overall, χ^2 (3, N = 98) = 2.38, p = .50, and χ^2 (3, N = 98) = 2.60, p = .46, respectively. Table V.3 shows the coefficients, associated Wald statistics, and significance levels for each of the predictors in the model.

Finally, due to the significant effects of case order on BPD diagnoses, the logistic regressions were repeated adding case order to the IMI scales as predictors of over- and under-diagnostic bias. Neither model was significant, χ^2 (3, N = 98) = 6.26, p = .10, and χ^2 (3, N = 98) = 4.59, p = .20, respectively, and for both models, case order was not found to be a significant predictor. Table V.4 shows the coefficients, associated Wald statistics, and significance levels for each of the predictors in the model.

Chapter 4

DISCUSSION

This study examined the role that clinicians' negative emotional reactions (NER) play in assigning the diagnosis of Borderline Personality Disorder (BPD). This disorder was selected as the main focus of the current investigation due to the well-documented prevalence of negative attitudes towards patients diagnosed with this condition (e.g., Gallop et al., 1989; Lenzenweger & Cicchetti, 2005; Markham, 2003; McIntyre & Schwartz, 1998; Nehls, 1998; Reiser & Levenson, 1984) and concerns that BPD may be misused and overdiagnosed. Since NER is strongly associated with a diagnosis of BPD, it was hypothesized that clinicians who report high levels of NER in response to a case would assign BPD more frequently and would be more likely to assign BPD without a sufficient number of DSM-IV symptoms to justify this diagnosis. Moreover, it was hypothesized that those reporting high NER would view other aspects of the case more negatively, such as the severity, prognosis, and the likelihood that the patient would benefit from treatment. Secondary hypotheses concerned the influence of clinician variables and order of symptom and diagnostic ratings.

Overall, support for the primary hypothesis was moderate, with some support for the influence of NER (i.e., generally at least one of the two IMI scales) on a diagnosis of BPD (versus other diagnoses), BPD representativeness ratings, BPD symptom ratings,

and ratings of prognosis and treatment response. However, predictions concerning the role of clinician variables (i.e., years of experience, percent of time spent in direct patient contact) and the influence of order of rating the symptoms and diagnoses received little support and, in some cases, the results were opposite of the prediction. In addition, gender of participants and case order had significant effects that had not been hypothesized.

The following discussion will address the hypotheses and findings as well as the role that other factors, such as symptom ratings and the order of the materials, played in the participants' responses to the cases. For each hypothesis, the results of both cases will be discussed and important differences between the cases will be highlighted. The relationship between the current findings and previous research will also be emphasized. Following the discussion of the hypotheses, the implications of this study will be addressed, with a specific focus on the assessment of BPD and training of future clinicians. The chapter will conclude with a discussion of the limitations of the study and implications for future research.

Relationship between NER and BPD Diagnoses and Ratings

The primary hypothesis that clinicians reporting higher levels of NER to the cases would be more likely to diagnose BPD, would assign higher BPD representativeness ratings, rate the case as more severe, the prognosis poorer, and the patient less likely to benefit from treatment received moderate support, although results were not entirely consistent across the two cases.

For Case One, there was a significant correlation between the Hostile scale and BPD representativeness ratings, and multiple regression showed higher Hostile ratings to

be a significant predictor of higher BPD ratings. However, further multiple regression analyses demonstrated that the Hostile scale influenced BPD representativeness ratings through its association with BPD symptom ratings (i.e., there was not a *direct* relationship between NER and BPD representativeness ratings). Therefore, participants who viewed the patient as more hostile considered the case as more representative of BPD, however, much of the variance was due their perception that there were more severe BPD symptoms in the case. This finding suggests that NER has a broad effect and influences the assessment of the presence and severity of BPD symptoms in the case, which are strong predictors of the degree to which the case is seen as representative of BPD.

Contrary to the hypothesis, however, higher levels of NER were not associated with more frequent diagnoses of BPD. Logistic regression revealed that although the Dominant scale was a significant predictor, unexpectedly, participants who viewed the patient as more dominant were actually *less* likely to assign the BPD diagnosis. However, this finding should be interpreted with caution because the Dominant scale was not significant as an individual predictor in the model. In addition, this finding may be partly explained by the relationship of the Dominant scale to Narcissistic Personality Disorder, as discussed below.

Results were also examined for Narcissistic Personality Disorder (NPD) because it was the second most frequently assigned diagnosis for the case and to help understand the effects of NER on differential diagnosis. NER, as measured by the Dominant and Hostile scales, was correlated with representativeness ratings, and median split analyses revealed that individuals rating higher on either IMI scale saw the case as more

narcissistic than participants who rated the IMI scales lower. However, regression analysis demonstrated that only the Dominant scale was a unique predictor, with higher scores predicting increased NPD representativeness ratings. One explanation for this finding is that a number of items on the Dominant scale (e.g., "she thinks other people find her interesting, amusing, fascinating, and witty;" "she wants me to put her on a pedestal;" "she makes me feel bossed around") are closely related to the DSM-IV criteria for NPD. Also, participants categorized as high on the Hostile scale assigned fewer diagnoses of NPD than participants categorized as low on the scale, although neither IMI scale was a significant predictor of a diagnosis of NPD. Overall, participants who viewed the patient as more dominant viewed the patient as more representative of NPD, which may help explain the finding that higher Dominant scale scores predicted fewer diagnoses of BPD. Moreover, the difference in NPD diagnoses between those categorized as high and low on the Hostile scale might relate to the Hostile scale predicting higher BPD symptom ratings.

For Case Two, there were significant correlations between both IMI scales and BPD diagnoses, representativeness ratings, and symptom ratings. Median split analyses revealed that participants classified as "high" on the Dominant scale assigned more BPD diagnoses whereas those classified as "high" on either scale assigned higher BPD representativeness ratings than those classified as "low." The regression analyses showed that higher ratings on the Dominant scale predicted BPD diagnosis and higher BPD representativeness ratings, even when controlling for BPD symptom ratings. These findings strongly suggest that NER (as measured by the IMI Dominant scale) was influential in determining participants' diagnosis of BPD and BPD representativeness

ratings for this case. Like Case One, Hostile scale scores were predictive of BPD symptom ratings suggesting that the perception of hostility affected the evaluation of BPD symptoms, which were then strong predictors of BPD representativeness ratings. One interesting finding with this case was that low scores on the Dominant scale were strong predictors of a diagnosis of Dependent Personality Disorder (DPD), both in the regression and median split analyses. For these clinicians, the patient's strong attachment to her boyfriend may have been interpreted as motivated by dependency needs rather than fears of being abandoned. As a result, these clinicians felt that the behavior was less dominant (i.e., manipulative) and more needy.

Together, results across both cases suggest moderate support for the role of NER in the diagnostic decision-making process related to BPD. NER, specifically clinicians' view that the patient exhibited hostility, affected their evaluation of BPD symptoms, with higher BPD symptom ratings when the patient was viewed as more hostile for both cases, and higher BPD symptom ratings were a strong predictor of BPD ratings. In addition, NER, primarily clinicians' perception of the patient as dominant, affected the differential diagnosis of BPD versus other disorders. Clinicians who viewed the patient as more dominant viewed the first case as more representative of NPD and the second case as more representative of BPD. Both BPD and NPD are cluster B (i.e., dramatic, emotional, erratic cluster) personality disorders characterized by externalizing behaviors, including the manipulative and exploitative behaviors described by the Dominant scale. These findings are consistent with previous studies demonstrating that clinicians view individuals with BPD as manipulative (e.g., Gallop et al., 1989; Lenzenweger & Cicchetti, 2005; Markham, 2003; McIntyre & Schwartz, 1998; Nehls, 1998; Reiser &

Levenson, 1984). In contrast, clinicians who saw the individual in Case Two as less dominant were more likely to diagnose Dependent Personality Disorder (DPD), a cluster C disorder characterized by submissive behaviors.

For the purposes of this study, NER was broadly defined by the IMI Hostile and Dominant scales. However, the two IMI scales differed in their relationship to diagnoses, representativeness ratings, and symptom ratings. One explanation for these differences may be found in the literature on the interpersonal circumplex and interpersonal theory (i.e., Kiesler, 1986). According to the principle of complementarity, hostility pulls for hostile responses (same end of the hostile-friendly dimension) whereas dominance pulls for submission (opposite end of the dominance-submission dimension). Based on the theory, it would be expected that hostility and dominance would differentially affect responses to the patients in the cases. Those who perceived the patients as more hostile would be expected to respond with more hostility (i.e., assign more BPD diagnoses and lower ratings of prognosis and response to treatment) whereas those who saw the patients as more dominant would respond in a more deferential way. Although results differed for the two scales, higher Dominant ratings did not result in less harsh ratings in the study.

Although the cases were constructed to be similar with regard to the number of BPD symptoms, there were considerable differences in participants' responses to the two cases. Clinicians rated Case One as somewhat more representative of BPD than Case Two (4.7 *vs.* 4.5) and assigned a diagnosis of BPD to the case more frequently (54.1% *vs.* 34.7%). For Case Two, clinicians appeared to regard the case as more indicative of Dependent Personality Disorder than BPD. Also, although both cases were designed to elicit equivalent levels of NER, Case One appeared to elicit stronger NER from

clinicians. There are several reasons that the findings might have differed between the cases.

One reason for the differences in participants' response to the cases was the difference in the symptoms included in the cases. Both cases were developed to exhibit two symptoms of BPD and to be subthreshold for the diagnosis. However, the cases contained different symptoms; Case One included unstable relationships and inappropriate, intense anger, and Case Two included affective instability and feelings of abandonment. Burgmer, Jessen, and Freyberger (2000) found that "a pattern of unstable and intense interpersonal relationships" was rated by clinicians as the most important symptom of BPD, followed closely by affective instability and inappropriate, intense, anger. Studies by Grilo and colleagues (Grilo, Becker, Anez & McGlashan, 2004; Grilo et al., 2002) also found that "a pattern of unstable and intense interpersonal relationships" predicted a diagnosis of BPD at a higher rate than the other symptoms included in the two cases in the present study. Therefore, Case One would be expected to receive more BPD diagnoses and higher BPD representativeness ratings than Case Two. Symptom ratings demonstrated that both cases exhibited the two symptoms that were intended to be present; however, Case One was also seen as demonstrating affective instability and Case Two was also seen as demonstrating unstable relationships (and symptoms from other diagnoses were rated as present in both cases). Therefore, a primary difference between BPD symptoms in the cases was the presence of intense anger in Case One and efforts to avoid abandonment in Case Two, which would also help explain the higher rate of BPD diagnoses in Case One and Dependent PD diagnoses in Case Two.

Another difference between the cases was that the symptomatic behaviors in Case One directly impact another person (i.e., displays of anger, devaluation of the therapist), whereas those in Case Two involved more internal states that primarily impact the patient (i.e., fears of abandonment, affective instability). Thus, clinicians responding to the cases might have been more sympathetic towards a patient whose symptoms did not cause suffering to another person and therefore assigned lower NER ratings. Additionally, the research conducted by Markam and Trower (2003) suggested that clinicians have more negative responses to noxious behaviors that they see as within the patient's control as opposed to outside of it. It is possible that the respondents believed that it is much easier to control *behaviors* (i.e., displays of anger, devaluation of the therapist) than *feelings* (i.e., fears of abandonment, affective instability) and judged the patient accordingly.

Ratings of Severity, Prognosis and Treatment Benefit

The hypothesis that those participants scoring high in NER would rate the case as more severe, assign lower prognosis ratings, and view the patient as being less likely to benefit from treatment received moderate support. For Case One, although both IMI scales were correlated with severity, neither scale was a significant predictor of severity ratings. However, high scores on the Hostile scale were shown to predict lower prognosis ratings and lower ratings that the patient would benefit from treatment, whereas the Dominant scale was not found be a significant predictor. For Case Two, higher scores on the Dominant scale predicted higher severity ratings and higher scores on the Hostile scale predicted lower ratings of the likelihood of benefiting from treatment.

Although the results differed between the cases, taken together these findings suggest that NER can influence clinicians' assessment of the patient's prognosis and

predictions of likely treatment effectiveness. These findings add to the previous literature that demonstrated that clinicians see patients with BPD as having a poor prognosis and being poor candidates for therapy (e.g., Hillman et al., 1997; McIntyre & Schwartz, 1998). The present results suggest that NER associated with the BPD symptoms may be a factor in clinicians' negative expectations of outcome for these patients.

Participant Characteristics

Years of Experience. The hypothesis that clinicians with fewer years of experience would be more susceptible to experiencing NER, and therefore assign higher BPD diagnostic and representativeness ratings, was not supported by this study. This finding is in contrast to those of Becker and Lamb (1994) who found that younger (i.e., less experienced) clinicians tended to assign BPD more readily in their vignette study. Morey and Ochoa (1989) also showed that clinicians with fewer years of experience were more likely to diagnose and even over-diagnose BPD. However, it is important to note that their finding was not replicated by Blashfield and Herkov (1996). Also, Garb's (1989) review of the literature on clinical decision making found no difference in accuracy of diagnoses between less and more experienced clinicians. In fact, the only difference found was that more experienced clinicians were found to be more accurate in their *confidence* ratings (i.e., they knew when they were right or wrong). Finally, years of experience was not found to be related to ratings of severity or predicted treatment benefit, although more experience was associated with lower ratings of prognosis for Case Two, perhaps as a consequence of more personal experience with patients with personality disorders.

Percent of Patient Contact. In contrast to the hypothesis and previous findings by Becker and Lamb (1994), participants with higher levels of patient contact were not more likely to assign a diagnosis of BPD, or report higher BPD representativeness ratings, than those having lower patient contact. Additionally, higher patient contact did not lead to increased ratings of severity or lower prognosis ratings for the case, although it was associated with lower ratings of the likelihood that the patient would benefit from treatment in Case Two. Participants with higher amounts of patient contact also reported higher Hostile ratings for Case Two; however, this was not associated with increased BPD ratings nor was this pattern seen in Case One. Since there was adequate variation in percent of direct patient contact among participants, it does not appear that the lack of support for the hypotheses was related to an under-representation of participants with high patient contact or a restricted range on this variable. One explanation for inconsistency with Becker and Lamb (1994) is that their participants included psychiatrists, social workers, and psychologists whereas the current study was composed entirely of psychologists. It is possible that important differences between these professional groups might have led to the findings in the Becker and Lamb study.

Gender. One interesting finding from the current study that was not part of the hypotheses was that female clinicians assigned a diagnosis of BPD more frequently than did male clinicians. Also, a within-subject comparison of participants' symptom ratings and diagnoses showed that female clinicians were more likely than male clinicians to over-diagnose BPD relative to their symptom ratings for Case One. This result is consistent with Morey and Ochoa's (1989) findings, although it was not replicated by Blashfield and Herkov (1996) nor was it found in Becker and Lamb's (1994) vignette

study. In the present study, this difference was not related to different NER ratings between men and women, so other explanations should be considered. For instance, it is possible that male clinicians were more influenced by the concern that women receive the diagnosis of BPD more often than men (APA, 1994). Men may have avoided diagnosing BPD for the women portrayed in the cases so as not to be perceived as demonstrating gender bias, whereas women may have been less concerned with appearing to be biased against their own gender. It should be noted that there were no gender differences on other variables, such as BPD symptom ratings, BPD representativeness ratings, and ratings of severity, prognosis, and treatment response.

A similar pattern of majority clinicians avoiding bias was found by Trierwieler et al. (2000) with respect to the diagnosis of schizophrenia in African-American patients. In that study, Caucasian clinicians (but not African-American clinicians) were *less* likely to assign a diagnosis of schizophrenia to an African-American patient than a Caucasian patient (the opposite of previous research findings). The researchers suggested that the majority clinicians may have avoided diagnosing schizophrenia in the African American patients due to concerns about race bias. Thus, it is possible that male clinicians in the current study might have been similarly concerned with gender bias.

Order Effects

Order of Diagnostic and Symptom Ratings. The order of symptom ratings and diagnostic ratings was counterbalanced to control for order effects. In addition, like the Crosby and Sprock (2004) study, order of the ratings was used to simulate the two decision-making methods identified by Blashfield and Herkov (1996): the prototype method, in which clinicians assign a diagnosis based on a comparison to an exemplar,

and the DSM approach, in which diagnostic criteria are carefully evaluated before assigning a diagnosis. Although Blashfield and Herkov (1996) suggested that neither approach is superior to the other, previous studies have suggested that a prototype approach might be associated with diagnostic bias compared to more structured evaluations of criteria (Hall, 2002), and it was hypothesized that the prototype approach would be associated with increased diagnoses of BPD and BPD representativeness ratings in this study.

The hypothesis that clinicians relying on a prototype approach would engage in more diagnostic bias was not supported in the current study. There was a non-significant trend for more BPD diagnoses and higher BPD representativeness ratings to be assigned when symptoms were evaluated first for Case One and no significant effects of order of diagnostic and symptom ratings for Case Two. In interpreting this finding, it is important to understand that the order of symptom and diagnostic ratings does not guarantee that participants used either the prototype or DSM-IV approach. In fact, clinicians asked to rate the symptoms first might have already formed a diagnostic impression while reading the case (i.e., a prototype approach) . Additionally, clinicians rating the symptoms first might have recognized that the patient met a number of core symptoms of BPD, increasing their likelihood of diagnosing BPD. It is important to note that all nine symptoms of BPD were present in the list of symptoms whereas only one or two symptoms of the other disorders were listed. Although this might have inadvertently influenced clinician ratings for the case, the BPD symptoms were randomly distributed within a list 42 total symptoms making it less likely that this significantly impacted the ratings.

Case Order. Although order of presentation of case was counterbalanced to control for order as a potential confound, the significant effects of case order on BPD diagnoses and ratings were unexpected. This effect was especially apparent for Case Two, as clinicians who received this case first were significantly more likely to assign a diagnosis of BPD and assign higher BPD representativeness ratings than those who received the case after reading Case One. Additionally, three BPD symptoms, including one of the intended symptoms, and NER (Dominant and Hostile scale) were rated higher and prognosis was rated lower for Case Two when it was presented first. One explanation for these findings is that Case One was generally seen as more representative of BPD based on the BPD diagnoses assigned (54.1% *vs.* 34.7%) and BPD representativeness ratings (4.7 *vs.* 4.5). As a result, when Case Two was viewed after Case One, clinicians may have viewed the case as less representative of BPD because they just viewed a more prototypical example of BPD in Case One. Further evidence for contrast effects are that "abandonment" was rated significantly lower for Case One when it was presented second whereas "anger" and "unstable relationships" were rated lower for Case Two when that case was presented second (i.e., participants rated the symptoms lower after viewing more extreme examples of the symptoms in the other case). In a review of the medical decision-making literature, both Hall (2002) and Elstein and Schwarz (2002) pointed out that order and contrast effects may be a common source of error in medical decision making.

Within-Subject Diagnostic Bias

The final hypothesis examined within-subject diagnostic bias by comparing a "criterion diagnosis" (determined by their BPD symptom ratings) with the assigned

diagnosis of BPD. Consistent with the primary hypothesis that NER would lead to increased diagnoses of BPD (i.e., over-diagnosis since the cases were subthreshold), it was hypothesized that participants scoring high in NER would over-diagnose BPD relative to their symptom ratings (i.e., assign a diagnosis of BPD without sufficient justification based on their symptom ratings). This hypothesis was not supported. NER was not associated with over- or under-diagnostic bias for either case. One explanation for this finding may be that NER, as defined by this study, was not the same as the "stress" referred to in the decision-making literature. For instance, studies demonstrating the negative effect of stress on decision-making (e.g., Edland, 1993; Edland & Svenson, 1993; Orasanu, 1997) have commonly used time pressure to induce stress. Additionally, Cumming and Harris (2001) showed that the effects of anxiety on decision making depended on the nature of the task assigned, with anxiety leading to errors more in simple tasks as opposed to complex tasks. Also, the decision-making literature suggests that negative affect may actually lead to more accurate and careful diagnostic decisions (Jundt & Hinsz, 2002). However, it is important to note the primary hypothesis that NER would increase BPD diagnoses and representativeness ratings received moderate support, and the fact that the cases were subthreshold suggests that the BPD diagnoses assigned to the cases represented over-diagnosis.

The finding that clinicians show disagreement between their "criterion diagnosis" based on symptom ratings and assigned diagnoses was similar to previous studies (i.e., Blashfield & Herkov, 1996; Crosby & Sprock, 2004; Morey & Ochoa, 1989). The current study found that half of the participants in Case One demonstrated inconsistency, primarily over-diagnosis, whereas one-third of the participants exhibited inconsistency

for Case Two, mostly under-diagnosis. However, NER does not appear to increase the frequency of this form of bias.

The last part of the hypothesis was that the order of rating symptoms and diagnoses would be a significant predictor of over-diagnostic bias, with those asked to assign a diagnosis before rating symptoms (i.e., prototype approach) being more likely to over-diagnose BPD relative to their symptom ratings than those asked to rate symptoms before assigning a diagnosis (i.e., DSM-IV) approach. Although the order of rating symptoms and diagnoses was a significant predictor of over- and under-diagnostic bias for Case One, results were in the direction opposite of prediction. Participants who received the symptom ratings first in Case One were more likely to *over-diagnose* BPD and those who received the diagnostic ratings first were more likely to *under-diagnose* BPD. These findings also diverge from those of Crosby and Sprock (2004). However, as discussed previously, order of the ratings does not guarantee that participants used that approach to decision making, and rating symptoms first resulted in more diagnoses of BPD than assigning a diagnosis first (the difference was not statistically significant). Although higher frequency of BPD diagnoses is not equivalent to within-subject over-diagnosis of BPD, the two are related as within-subject over-diagnosis can only occur if a diagnosis of BPD is assigned. Thus, the proposed explanation for the increased diagnoses of BPD when symptoms were rated first (i.e., participants became more aware of the BPD symptoms in the case) is probably relevant to explaining these findings.

Implications

Overall, this study provided moderate support for the hypothesis that NER would influence the way clinicians use the BPD diagnosis, and suggest that the relationship is

due to the effect of NER on the perception of BPD symptoms as well as consideration of the representativeness of BPD and alternative diagnoses (i.e., NPD and DPD for these two cases). The current findings add to the research on the stigma and negative emotional reactions of clinicians to a diagnosis of BPD (e.g., Gallop et al., 1989; Lenzenweger & Cicchetti, 2005; Markham, 2003; McIntyre & Schwartz, 1998; Nehls, 1998; Reiser & Levenson, 1984). Although past studies focused on NER as a consequence of a BPD diagnosis, the contribution of NER in *formulating* the diagnosis of BPD was universally ignored. The findings of this study suggest that it is possible that negative reactions towards patients contribute to the diagnosis of BPD rather than only being a consequence of it. Thus, future research on the stigma associated with BPD might devote more time to critically examining the factors that lead to a diagnosis of BPD in the first place.

In addition, the results of this study carry implications for the training of future clinicians. According to a study conducted by Cleary, Siegfried, and Walter (2002), 95% of a sample of 199 staff members working at a mental health outpatient clinic reported that they would desire more training in the proper assessment and management of BPD patients due to the difficulties posed by this population. Results of the present study suggest that training to help clinicians reduce their negative reactions to patients with BPD (and perhaps personality disorders in general since NER also predicted NPD ratings) might improve differential diagnosis and subsequently, treatment. In fact, a study conducted by Krawitz (2004) showed that a two-day workshop (emphasizing etiology and current concepts of BPD) was effective in reducing clinicians' negative attitudes and increasing their willingness to work with BPD patients. Perhaps one

element that could be included in such a training program is the importance of not allowing one's emotional reaction to overly influence judgments in the diagnosis of BPD. Counter-transference reactions are an inevitable part of working with patients with personality disorders and are understandably used as an informal aid to diagnosis (Westen, 1997). However, using sound clinical judgment when diagnosing BPD is especially important given the findings that patients who are diagnosed with BPD could experience negative consequences such as having their symptoms viewed as "attention-seeking" and manipulative (i.e., Gallop et al., 1989; Markam & Trower, 2003), their prognosis poorer, and being less likely to benefit from treatment (Gallop et al., 1989; Hillman et al., 1997). Finally, this study also confirmed the finding by Morey and Ochoa (1989) that women are more likely than men to over-diagnose BPD, suggesting the need for future research to examine the factors contributing to gender differences in diagnosis of BPD.

Limitations

In reviewing the results for a BPD diagnosis, it is important to keep in mind that clinicians were forced to choose an Axis I and II diagnosis for the case. In other words, clinicians were not given the option of choosing "no diagnosis" or "diagnosis deferred," although they could opt to leave the item blank. Although confidence ratings were relatively high for Axis II diagnosis for both cases, it does not necessarily follow that clinicians would actually *assign* the diagnosis in a real word setting. When given the opportunity to provide a comment at the end of the study, several of the participants expressed concerns about this issue, with one stating, "I would prefer to have a category that would include 'insufficient information to assess at this time." In addition, despite

the high overlap and comorbidity of personality disorder diagnoses (Clark, Watson & Reynolds, 1995), participants could only select one axis II diagnosis, which may have affected the results (i.e., diagnosis of BPD versus NPD or DPD). Thus, it is important to keep in mind that there are differences between the diagnoses and diagnostic ratings assigned in this study and diagnoses assigned in a real clinical setting. On the other hand, the diagnostic representativeness ratings allowed clinicians to use a dimensional approach to diagnosis, and the majority of the significant findings were for the dimensional ratings rather than for the categorical diagnoses.

Another limitation of the current study is the use of the Impact Message Inventory (IMI) as a measure of negative emotional reaction towards the cases. This measure has frequently been used to measure the emotional impacts that patients have on their therapists (e.g., Stephens et al., 1987; Swaney & Stone, 1990; Simon et al. 1992), but it was designed to be used during live encounters, although McIntyre and Schwartz (1998) used the IMI to assess participants' responses to taped interviews. Thus, a major modification to the IMI for this study was that participants were to simply *imagine* they were in the presence of the person depicted in the vignette, which may have limited the validity of the results. Additionally, because the full 90-item IMI was impractical for use in this study, only the Dominant and Hostile scales were used. Although the use of these scales was justified by McIntyre and Schwartz's (1998) findings concerning mental health professionals' responses to patients with BPD, this limited the range of emotional responses to the cases that were examined. Related, the present study considered NER broadly and did not formulate specific hypotheses for the two IMI scales, yet the scales

differentially predicted BPD diagnoses, BPD representativeness ratings, and BPD symptom ratings.

Another limitation of the study is that the results might have been due to the unique characteristics of the two cases, such as the specific BPD symptoms included and the behaviors used to elicit NER. It is possible that different results would have been obtained if other BPD symptoms were included in the cases. For example, the BPD symptoms in the cases were selected to be moderately associated with a diagnosis of BPD because the use of heuristics increases with uncertainty (Hall, 2002). Neither case included suicidal ideation or self-injurious behavior (and this was confirmed by the symptom ratings) because of findings that these symptoms are highly predictive of a diagnosis of BPD (Grilo et al., 2004; Grilo et al., 2002). On the other hand, these symptoms have been shown to predict over-diagnosis of BPD (i.e., Blashfield & Herkov, 1996; Morey and Ochoa, 1989).

A related limitation is the use of vignettes (Hughes & Huby, 2002). First, because vignettes usually focus on several key variables, they do not present the full complexity of an actual case nor the opportunity to interact with a real patient, limiting generalizeability to the real world. Next, vignettes have been criticized because the information in them is not easily retained compared to observed behavior, although participants in the present study were allowed to review the cases while completing the ratings. Finally, vignettes are limited in their ability to stimulate genuine emotional responses from participants. In this study, NER ratings were only moderate for Case One and low to moderate for Case Two, whereas interaction with actual patients exhibiting these behaviors might be expected to elicit much stronger emotional responses. Although

these limitations may have influenced the results of the study, vignettes have been widely used to examine diagnostic decision making because they allow for control of the variables of interest.

Additionally, there are disadvantages to using the Internet to collect data including the potential for multiple submissions or data sabotage (e.g., participants who were not qualified to participate in the study). However, only two instances of duplicate data were found and were removed from the database, and psychologists who were invited to participate were assigned a password that they entered when they signed onto the website (although this password was not unique for each participant due to concerns over confidentiality). Another disadvantage of using the Internet is lack of control over the data collection environment (Whitley, 2002); participants may have enlisted help or consulted the DSM-IV, even though they were instructed not to utilize outside resources. However, much of the research on diagnostic decision making in psychopathology has used mail surveys (e.g., Blashfield & Herkov, 1996; Funtowicz & Widiger, 1999; Morey & Ochoa, 1989; Samuel & Widiger, 2006; Westen, Shedler, Durrett, Glass, & Martens, 2003) which have the same disadvantage. Finally, use of the Internet might lead to non-random selection of participants (Birnbaum, 2004) and participation biases (e.g., lower rates of participation from older psychologists and those with less technological expertise). However, the characteristics of this sample do not suggest that younger clinicians were over-represented in the population. On the other hand, the Internet has several major advantages as a research tool, such as reducing data-entry errors by allowing participants to automatically submit their data to a database and reducing the cost of conducting a research study (Birnbaum, 2004; Michalak & Szabo, 1998).

A final limitation of this study is the relatively small sample size and low response rate, which might limit the generalizability of the results, particularly in light of the selection biases discussed above. The 8% response rate found in this study is markedly smaller than the 10-20% response rate found in other studies using clinicians as participants (e.g., Crosby & Sprock, 2004; Samuel & Widiger, 2004; Sprock, 2003). One reason for this may relate to the use of the Internet for data collection, rather than mail surveys, particularly the use of email invitations which can easily be read and forgotten or deleted. Also, because participants were predominantly Caucasian and very few ethnic minorities participated in the study, it is difficult to generalize the results to the broader population of practicing psychologists. However, the demographic characteristics of the sample approximate those of the most recent APA membership profile (see American Psychological Association website: http://research.apa.org/2002membershipt1.pdf) with respect to gender and age (data on ethnicity was not reported by approximately one-fourth of APA members).

Directions for Future Research

Given the limitations relating to the use of the IMI as the measure of NER, future research might focus on developing instruments to gauge the emotional reaction of clinicians towards their patients. Since a diagnosis of BPD is so strongly associated with negative attitudes from clinicians (e.g., Gallop et al., 1989; Markham, 2003; McIntyre & Schwartz, 1998; Nehls, 1998; Reiser & Levenson, 1984) and data suggest that clinicians desire increased training on how their own emotional reactions influence their treatment of BPD patients (Cleary et al., 2002), the development of an instrument that specifically measures common negative attitudes towards BPD patients (e.g., manipulative, attention-

seeking, etc.) is certainly warranted. Such a measure would be ideal for use in future research addressing how NER contributes to the formulation of a diagnosis of BPD and expectations for response to treatment. Future research might also investigate the influence of specific types of NER and their effect on diagnosis and treatment.

Another direction for future research would be to examine NER and BPD diagnoses for different symptoms of BPD. Given the different results for the two cases in the current study, a more thorough investigation of the relationship of NER to the remaining BPD symptoms is certainly warranted. However, if such a study included recurrent self-injurious behaviors or suicidal gestures in the vignette, it is unlikely that enough participants would assign diagnoses *other* than BPD and this would limit an investigation into the role of NER.

Since vignettes have limited utility in eliciting genuine emotional reactions from participants, another avenue for future research would be to conduct similar studies using more vivid and realistic depictions of patients. For instance, the use of a photograph, audiotape, or videotape of real or simulated interviews would allow more opportunities to elicit NER from participants and increase the external validity of the findings. In such a study, participants would actually get to hear and observe important characteristics of the patient, such as a sarcastic tone of voice or a condescending facial expression. The increasing use of the Internet for data collection makes the possibility of including such stimuli in future research highly feasible. Although this would certainly lead to a more rich clinical presentation, researchers would need to carefully select and engineer the stimuli used so as not to include unanticipated variables (Birnbaum, 2004).

A related possibility would be to use a methodology similar to Morey and Ochoa (1989) and Blashfield and Herkov (1996). As discussed previously, these studies asked clinicians to identify an actual patient, report their diagnoses and provide symptom ratings. Clinicians could be asked to identify a difficult patient, as in the Bongar et al. (1991) study and then provide a diagnosis, rate symptoms, and rate NER. It is likely that asking clinicians to rate actual patients would elicit a higher level of NER than was possible in the current study. Although the relationship of NER to over- and under-diagnostic bias could certainly be investigated, it would not address how NER contributes to a diagnostic formulation (i.e., clinicians would likely have already assigned a diagnosis to a patient they were familiar with). An alternative approach would be to ask clinicians to provide their diagnostic impressions, symptom ratings, and NER ratings after an intake assessment or one-time crisis intervention assessment in order to examine NER during the process of formulation of a diagnosis. In this situation, clinicians would have less information to support their diagnosis and might be more vulnerable to the potential sources of diagnostic bias.

Future research might also address the finding that female clinicians are more likely to over-diagnose BPD relative to male clinicians. Because Morey and Ochoa (1989) also found women to over-diagnose BPD, it is important to begin addressing the reasons underlying this pattern. As discussed previously, male clinicians might be more sensitive to the potential for gender bias (i.e., assigning BPD to a female patient). Thus, future studies might ask participants to rate the importance of patient demographic characteristics in assigning their diagnosis, such as the patient's age or gender. Also, clinicians could be directly asked about their concerns of gender bias when assigning a

diagnosis of BPD in a separate study in order to determine if differences between male and female clinicians exist on this variable. If male clinicians do not appear to be especially sensitive to the potential for gender bias, then other explanations, such as the way in which female clinicians perceive pathology in other women, should be explored.

Another direction for future research would be to determine if NER affects diagnostic decisions for men who exhibit features of BPD. Since Becker and Lamb (1994) showed that male patients were less likely to receive a BPD diagnosis than female patients with the same symptoms, perhaps men would be more likely to receive a different diagnosis, such as Antisocial Personality Disorder or Narcissistic Personality Disorder, in response to heightened levels of NER.

A final direction for future research could be to investigate the role of NER in assigning other personality disorder diagnoses and Axis I diagnoses. Since NER was shown to have some impact on BPD diagnosis and representativeness ratings, as well as representativness ratings for Narcissistic Personality Disorder and Dependent Personality Disorder, it is possible that NER might influence other Axis II and even Axis I diagnoses. It would also be worthwhile to examine the interaction of Axis I and Axis II diagnoses. Clinicians who experience higher levels of NER might be more likely to assign a personality disorder diagnosis and less likely to assign an Axis I diagnosis, such as Major Depressive Disorder, because it is not associated with a negative emotional reaction (McIntyre & Schwartz, 1998). It would also be interesting to examine the relationship of NER to Axis I disorders that might be associated with stigma, such as Factitious Disorder, and to investigate the influence of NER on diagnosis and clinical decision making.

REFERENCES

REFERENCES

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, D. C.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., revised). Washington, D. C.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed). Washington, D. C.

Becker, D. & Lamb, S. (1994). Sex bias in the diagnosis of borderline personality disorder and posttraumatic stress disorder. *Professional Psychology: Research and Practice, 25(1),* 55-61.

Bergus, G. R., & Hamm, R. M. (1995). How physicians make medical decisions and why medical decision making can help. *Primary Care: Clinics in Office Practice, 22(2),* 167-180.

Birnbaum, M. (2004). Human research and data collection via the Internet. *Annual Review of Psychology, 55,* 803-832.

Blashfield, R.K., Sprock, J., Pinkston, K. & Hodgin, J. (1985). Exemplar prototypes of personality disorder diagnoses. *Comprehensive Psychiatry, 26,* 11-21.

Blashfield, R.K., Sprock, J., Haymaker, D. & Hodgin, J. (1989). The family resemblance hypothesis applied to psychiatric classification. *Journal of Nervous and Mental Disease, 177,* 492-497.

Blashfield, R., & Herkov, M. (1996). Investigating clinician adherence to diagnosis by criteria: A replication of Morey and Ochoa (1989). *Journal of Personality Disorders, 10* (3), 219-228.

Boggs, C. D., Morey, L. C., Skodol, A. E., Shea, M. T., Sanislow, C. A., Grilo, C. M., McGlashan, T. H., Zanarini, M. C., and Gunderson, J. G. (2005). Differential impairment as an indicator of sex bias in DSM-IV criteria for four personality disorders. *Psychological Assessment, 17(4),* 492-496.

Bongar, B., Markey, L. A., & Peterson, L. G. (1991). Views on the difficult and dreaded patient. A preliminary investigation. *Medical Psychotherapy: An International Journal, 4,* 9-16.

Briggs-Phillips, M. L. (2004). The impact of the stigma of borderline personality disorder diagnosis on treatment expectations of therapists. *Dissertation Abstracts International, 64(10B),* 5208.

Burgmer, M., Jessen, F., & Freyberger, J. (2000). Polythetic diagnostic approach to the borderline personality disorder: The valency of the single criterion in the concept of professional therapists. *Psychopathology, 33(3),* 119-124.

Carpenter, B.N., & Suhr, P. (1998, August). *Stress appraisal: Measurement and correlates.* Paper presented at the annual meeting of the American Psychological Association, Atlanta, GA.

Clark, L.A., Watson, D. & Reynolds, S. (1995). Diagnosis and classification. *Annual Review of Psychology, 46,* 121-153.

Cleary, M., Siegfried, N, & Walter, G. (2002). Experience, knowledge, and attitudes of mental health staff regarding clients with a borderline personality disorder. *International Journal of Mental Health Nursing, 11,* 186-191.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37-46.

Cohen J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159.

Crosby, J. & Sprock, J. (2004). The effect of patient sex and clinician sex and sex role in the diagnosis of Antisocial and Histrionic personality disorder. *Journal of Clinical Psychology, 60(6),* 1-22.

Cumming, S. R. & Harris, L. M. (2001). The impact of anxiety on the accuracy of diagnostic decision-making. *Stress and Health, 17,* 281-286.

Edland, A. (1993). The effects of time pressure on choices and judgments of candidates to a university program. In Svenson, O. & Maule, A. J. (Eds.) *Time Pressure and Stress in Human Judgment and Decision Making*, pp. 145-156. New York: Plenum Press.

Edland, A. & Svenson, O. (1993). Judgment and decision making under time pressure: Studies and findings. In Svenson, O. & Maule, A. J. (Eds.) *Time Pressure and Stress in Human Judgment and Decision Making*, pp. 27-38. New York: Plenum Press.

Elstein, A. S. & Schwarz, A. (2002). Clinical problem solving and diagnostic decision making: Selective review of the cognitive literature. *British Medical Journal, 324*, 729-732.

Ford, M., & Widiger, T. (1989). Sex bias in the diagnosis of histrionic and antisocial personality disorders. *Journal of Consulting and Clinical Psychology, 57* (2), 301-305.

Funtowicz, M. N. & Witiger, T. (1999). Sex bias in the diagnosis of personality disorders: An evaluation of the DSM-IV criteria. *Journal of Abnormal Psychology, 108(2),* 195-201.

Gallop, R., Lancee, W. J., & Garfinkel, P. (1989). How nursing staff respond to the label "borderline personality disorder." *Hospital and Community Psychiatry, 40(8),* 815-819.

Garb, H. N. (1989). Clinical judgment, clinical training, and professional experience. *Psychological Bulletin, 105(3),* 387-396.

Garb, H. N. (1997). Race bias, social class bias, and gender bias in clinical judgment. *Clinical Psychology: Science and Practice, 4(2),* 99-120.

Gilliland, B. E., & James, R. K. (1997). *Crisis Intervention Strategies (3rd Ed.).* Pacific Grove: Brooks/Cole Publishing Company.

Giacalone, R. C. (1997). A study of clinicians' attitudes and sex bias in the diagnosis of borderline personality disorder and posttraumatic stress disorder. *Dissertation Abstracts International, 57(12-B),* 7725.

Grilo, C. M., McGlashan, T. H., Morey, L. C., Gunderson, J. G., Skodol, A. E., Shea, M. T., Sanislow, C. A., Zanarini, M. C., Bender, D., Oldham, J. M., Dyck, I., & Stout, R. (2001). Internal consistency, intercriterion overlap and diagnostic efficency of criteria sets for DSM-IV schizotypal, borderline, avoidant, and obsessive-compulsive personality disorders. *Atca Psychiatrica Scandinavica, 104(4),* 264-272.

Grilo, C. M., Becker, D. F., Anez, L. M., & McGlashan, T. H. (2004). Diagnostic efficiency of DSM-IV criteria for borderline personality disorder: An evaluation in Hispanic men and women with substance use disorders. *Journal of Clinical and Consulting Psychology, 72(1),* 126-131.

Gruppen, L. D., Wolf, F. M., & Billi, J. E. (1991). Information gathering and integration as sources of error in diagnostic decision making. *Medical Decision Making, 11*, 233-239.

Gunderson, J. G., Zanarini, M. C., & Kisiel, C. L. (1995). Borderline Personality Disorder. In W. J. Livesley (ed.), *The DSM-IV Personality Disorders*. New York: Guilford Press, 141-157.

Hall, K. H. (2002). Reviewing intuitive decision-making and uncertainty: the implications for medical education. *Medical Education, 36*, 216-224.

Hillman, J. L., Stricker, G., & Zweig, R. A. (1997). Clinical psychologists' judgments of older adult patients with character pathology: Implications for practice. *Professional Psychology: Research and Practice, 28(2),* 179-183.

Hughes, R. & Huby M. (2002). The application of vignettes in social and nursing research. *Journal of Advanced Nursing, 37(4),* 382-386.

Jundt, D. K. & Hinsz, V. B. (2002). Influences of positive and negative affect on decisions involving judgmental biases. *Social Behavior and Personality, 30(1),* 45-52.

Kiesler, D. (1986). The 1982 interpersonal circle. In T. Millon & G. Klerman (Eds.), *Contemporary directions in psychopathology* (pp. 571-597). New York: Guilford Press.

Kiesler, D. J., & Schmidt, J.A. (1993). The Impact Message Inventory: Form IIA Octant Scale Version. Redwood City, CA: Mind Garden.

Krawitz, R. (2004). Borderline personality disorder: Attitudinal change following training. *Australian and New Zealand Journal of Psychiatry, 38(7),* 554-559.

Lenzenweger, M. F. & Cicchetti, D. (2005). Toward a developmental psychopathology approach to borderline personality disorder. *Development and Psychopathology, 17(4),* 893-898.

MacCallum, R.C., Zhang, S., Preacher, K.J., & Rucker, D.D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7,* 19-40.

Markham, D. (2003). Attitudes towards patients with a diagnosis of 'borderline personality disorder': Social rejection and dangerousness. *Journal of Mental Health, 12(6),* 595-612.

Markham, D. & Trower, P. (2003). Effects of the psychiatric label 'borderline personality disorder' on nursing staff's perceptions of causal attributions for challenging behaviors. *British Journal of Clinical Psychology, 42(3),* 243-256.

Maule, A. J. & Svenson, O. (1993). Theoretical and empirical approaches to behavioral decision making and their relation to time constraints. In Svenson, O. & Maule, A. J. (Eds.) *Time Pressure and Stress in Human Judgment and Decision Making,* (pp. 3-22). New York: Plenum Press.

McGlashan, T. H. (1987). Testing DSM-III symptoms criteria for schizotypal and borderline personality disorders. *Archives of General Psychiatry, 44(2),* 143-148.

McIntyre, S. M. & Schwartz, R. C. (1998). Therapists' differential countertransference reactions toward clients with major depression or borderline personality disorder. *Journal of Clinical Psychology, 54(7),* 923-931.

Michalak, E. E. & Szabo, A. (1998). Guidelines for Internet research: An update. European Psychologist, 3(1), 70-75.

Morey, L. C. & Ochoa, E. S. (1989). An investigation of adherence to diagnostic criteria: Clinical diagnosis of the DSM-III personality disorders. *Journal of Personality Disorders, 3(3),* 180 – 192.

Murray, D. M., & Fisher, J. D. (2002). The internet: A virtually untapped tool for research. *Journal of Technology in Human Services, 19(2,3),* 5-18.

Nehls, N. (1998). Borderline personality disorder: gender stereotypes, stigma, and limited system of care. *Issues in Mental Health Nursing, 19,* 97-112.

Ochoa, E. S., & Morey, L. C. (1990). Clinical diagnosis of borderline personality: examination of professional differences. *Professional Psychology: Research and Practice, 21(1),* 54-59.

Orasanu, J. (1997). Stress and naturalistic decision making: Strengthening the weakest links. In Flin, R., Salas, E., Strub, M., & Martin, L. (Eds.) *Decision Making Under Stress: Emerging Themes and Applications,* (pp. 43-66). Aldershot: Ashgate Publishing.

Reich, J. (1987). Sex distribution of DSM-III personality disorders in psychiatric outpatients. *American Journal of Psychiatry, 144(4),* 485-488.

Reiser, D. E. & Levenson, H. (1984). Abuses of the borderline diagnosis: A clinical problem with teaching opportunities. *American Journal of Psychiatry, 14(12),* 1528-1532.

Samuel, D. B. & Widiger, T. (2006). Clinicians' judgments of clinical utility: A comparison of the DSM-IV and Five-Factor Models. *Journal of Abnormal Psychology, 115(2),* 298-308.

Sansone, R. A., Songer, D. A., & Gaither, G. A. (2001). Diagnostic approaches to borderline personality and their relationship to self-harm behavior. *International Journal of Psychiatry in Clinical Practice, 5,* 273-277.

Sansone, R. A., Songer, D. A., & Miller, K. A. (2005). Childhood abuse, mental healthcare utilization, self-harm behavior, and multiple psychiatric diagnoses among inpatients with and without a borderline diagnosis. *Comprehensive Psychiatry, 46(2),* 117-120.

Simon, L., Gaul, R., Friedlander, M. L., & Heatherington, L. (1992). Client gender and sex role: Predictors of counselors impressions and expectations. *Journal of Counseling and Development, 71,* 48-52.

Stephens, R. S., Hokanson, J. E., & Welker, R. (1987). Responses to depressed interpersonal behavior: Mixed reactions in a helping role. *Journal of Personality and Social Psychology, 52,* 1274-1282.

Svenson, O. (1997). Differentiation and consolidation theory. In Flin, R., Salas, E., Strub, M., & Martin, L. (Eds.) *Decision Making Under Stress: Emerging Themes and Applications*, (pp. 301-314). Aldershot: Ashgate Publishing

Swaney, K. B., & Stone, G. L. (1990). Therapist awareness of covert reactions to client interpersonal behavior. *Journal of Social and Clinical Psychology, 9,* 375-389.

Taylor, S. (1995). Commentary on Borderline Personality Disorder. In W. J. Livesley (Ed.), *The DSM-IV Personality Disorders*, New York: Guliford Press, 1995, 165-172.

Trierweiler, S. J., Neighbors, H. W., Munday, C., Thompson, E. E., Binion, V. J. & Gomez, J. P. (2000). Clinician attributions associated with the diagnosis of schizophrenia in African American and Non-African American patients. *Journal of Consulting and Clinical Psychology, 68(1),* 171-175.

Warner, R. (1978). The diagnosis of antisocial and hysterical personality disorders. An example of sex bias. *Journal of Nervous and Mental Disease, 166,* 839-845.

Westen, D. (1997). Divergences between clinical and research methods for assessing personality disorders: Implications for research and the evolution of Axis II. *American Journal of Psychiatry, 154(7),* 895-903.

Westen, D., Shedler, J., Durrett, C., Glass, S., Martens, A. (2003). Personality diagnoses in adolescence: DSM-IV Axis II diagnoses and an empirically derived alternative. *American Journal of Psychiatry, 160(5),* 952-967.

Whitley, B. E. (2003). *Principles of Research in Behavioral Science.* Boston: McGraw-Hill.

Widiger, T. (1998). Invited essay: Sex biases in the diagnosis of personality disorders. *Journal of Personality Disorders, 12* (2), 95-118.

Widiger, T., & Spitzer, R. (1991). Sex bias in the diagnosis of personality disorders: Conceptual and methodological issues. *Clinical Psychology Review, 11,* 1-22.

Widiger, T. A. & Trull, T. J. (1993). Borderline and Narcissistic Personality Disorders. In Sutker, P. B. & Adams, H. E. (Eds.) *Comprehensive Handbook of Psychopathology,* 2nd Edition. New York: Plenum Press, 371-390.

Widiger, T. A., & Weissman, M. M. (1991). Epidemiology of Borderline Personality Disorder. *Hospital and Community Psychiatry, 42(10),* 1015-1020.

Zlotnick, C., Rothschild, L., & Zimmerman, M. (2002). The role of gender in the clinical presentation of patients with borderline personality disorder. *Journal of Personality Disorders, 16(3),* 277-283.

APPENDIXES

Appendix A

Diagnostic Criteria for Borderline Personality Disorder

A pervasive pattern of instability of interpersonal relationships, self-image, and affects, and marked impulsivity beginning in early adulthood and present in a variety of contexts, as indicated by five (or more) of the following:

1. frantic efforts to avoid real or imagined abandonment. Note: Do not include suicidal or self-mutilating behavior covered in Criterion 5.
2. a pattern of unstable and intense interpersonal relationships characterized by alternating between extremes of idealization and devaluation.
3. identity disturbance: markedly and persistently unstable self-image or sense of self.
4. impulsivity in at least two areas that are potentially self-damaging (e.g., spending, sex, substance abuse, reckless driving, binge eating). Note: Do not include suicidal or self-mutilating behavior covered in Criterion 5.
5. recurrent suicidal behavior, gestures, or threats, or self-mutilating behavior.
6. affective instability due to a marked reactivity of mood (e.g., intense episodic dysphoria, irritability, or anxiety usually lasting a few hours and only rarely more than a few days).
7. chronic feelings of emptiness
8. inappropriate, intense anger or difficulty controlling anger (e.g., frequent displays of temper, constant anger, recurrent physical fights).
9. transient, stress-related paranoid ideation or severe dissociative symptoms.

Appendix B

Case Vignettes

#1

Mrs. J., a 35 year-old female, presents to your office for an intake interview after failing to attend her first scheduled appointment one week ago and not calling to cancel. When the client arrives a few minutes late at your office, you hear her complaining loudly to the office staff that nearly all of the chairs are occupied in the waiting room and she does not have enough space to fill out her paperwork. Additionally, when you greet Mrs. J. in the waiting room she complains sarcastically that "with all that money you're charging me you should probably invest in some more chairs."

According to Mrs. J., her husband wants her to get help with her "anger problems." She states that her husband is tired of her frequent outbursts towards him. Mrs. J. stated that her husband threatened to leave her last week if she did not seek therapy after a rather heated argument about establishing a proper curfew for their 16 year-old daughter. Mrs. J. admitted that she has somewhat of a "temper" but maintained that "it's not as bad as everyone makes it out to be." When describing these events to you, you notice that Mrs. J. becomes very agreeable and complimentary towards you when you attempt to validate her frustration with her husband. Specifically, when you say "I see how that could be frustrating for you to deal with" Mrs. J. states, "Yeah, I know. You're so much better than those other therapists, you must have gone to a really good school." However, she quickly becomes angry towards you when you encourage her to consider alternative ways of coping with her husband's behavior, stating, "you

know, you're just like the rest of them, always saying that I'm the problem. I knew I shouldn't have even come here, you're all quacks!"

Mrs. J. indicated that she has been in therapy multiple times to help her manage her anger. She stated that she has terminated therapy each of these times due to her belief that the therapist did not want to help her or did not understand her situation. According to records obtained from her previous therapists, she has been noncompliant with treatment and exhibited a similar pattern of interaction with them. In addition, Mrs. J. reported that she has been prescribed various medications in the past to help her with these symptoms but admitted that she rarely takes the medication. She denied having any history of violent behavior.

#2

Ms. S. is a 28 year old female who presents fifteen minutes late at your clinic for her intake appointment. When asked about the reasons for being late, Ms. S. states "oh, I didn't even look at the time. I didn't think it would be a big deal." According to Ms. S, she was advised to seek therapy by her primary care physician due to his concern about her recurrent emotional distress associated with her boyfriend. She stated that she does not believe that her physician has made the appropriate recommendation, stating "I'm just tired of talking about this. I just want someone to fix me so my boyfriend will stay with me. I don't see how going through all this again is going to help anything. I think my doctor just wanted to get rid of me. I don't think he really likes working with female patients either. That's not ethical, is it? I think you and I should file a complaint against him."

Information from the referral states that Ms. S. frequently becomes depressed when her boyfriend comes home from work late or makes plans that do not involve her. Ms. S. explains that "when he doesn't come home when he should, it's just terrible. He should be more considerate of my time and know that I'm going to feel so anxious and terrible. Why can't he understand this?" When describing this pattern, Ms. S. begins sobbing loudly and states, "see how he makes me feel? It just isn't right." Ms. S. states that she feels this way when he is only five or ten minutes late from work and that her feelings are worse when he is out spending time with his friends. However, she states that her mood quickly improves when he returns home. When the topic shifts to a discussion of a more neutral topic (e.g., her work history), her mood quickly improves and she provides answers to all questions asked.

When asked for further detail about her behavior when her boyfriend does not return home as expected, Ms. S. admits that she often will call his workplace multiple times to find out if anyone knows where he is located. Additionally, she states "My doctor thought this may be causing the problems in my relationship, as if all this is my fault." Ms. S. states that she has experienced similar feelings in previous relationships but denied that her inability to tolerate being alone has led previous partners to terminate the relationship.

Ms. S. stated that she has sought psychotherapy services in the past but that her therapist was not especially helpful. She reports that "she just wasn't helping me and wanted me to do all the work. She wanted me to do all this stupid stuff outside of the session (e.g., mood monitoring). I hope you don't ask me to do that type of stuff. It was just silly. If she really knew what she was doing, she would have fixed me during the

appointment." Records from her previous treatment indicate that she attended sessions sporadically and that she was generally non-compliant with homework assignments.

Appendix C

Piloting

In order to determine the effectiveness of the two vignettes in eliciting NER reactions, the vignettes were presented to a group of doctoral students in clinical psychology at Indiana State University (ISU). They were asked to read the vignettes and respond to the selected items from the IMI. Based on the ratings and feedback from the group, the vignettes were revised to ensure that varying levels of NER would be elicited from the sample. In addition, a group of students were asked to complete the diagnostic and symptom ratings for the vignettes. Based on the results, the vignettes appeared to contain the intended symptoms and were seen as moderately representative of a diagnosis of BPD. Once the vignettes and the survey questions were finalized and placed on the Internet, a group of students were recruited to complete the entire survey. Information from that round of testing was used to identify points of confusion among the participants regarding the task, estimate the amount of time necessary to read the vignettes and to complete the ratings, and to locate any technical problems that may interfere with the data collection process. Feedback from this group was incorporated into the study. Finally, a series of data was entered for each of the four web pages and checked against hard copies of the data to ensure that the information was properly coded and transferred to the database.

Appendix D

Diagnostic and Symptom Ratings

Rate the applicability of the following diagnoses for this patient using the following scale:

1 2 3 4 5 6 7

Not at all Representative Very Representative

Axis I Disorders

___ Major Depressive Disorder
___ Intermittent Explosive Disorder
___ Bipolar I Disorder
___ Dysthymic Disorder
___ Adjustment Disorder
___ Bipolar II Disorder

Axis II (Personality Disorders)

___ Paranoid PD
___ Schizoid PD
___ Schizotypal PD
___ Antisocial PD
___ Borderline PD
___ Histrionic PD
___ Narcissistic PD
___ Avoidant PD
___ Dependent PD
___ Obsessive-Compulsive PD
___ Passive-Aggressive PD
___ Depressive PD

Choose the <u>one</u> Axis I diagnosis most representative of the above case

1. Major Depressive Disorder
2. Intermittent Explosive Disorder
3. Bipolar I Disorder
4. Dysthymic Disorder
5. Adjustment Disorder
6. Bipolar II Disorder

Choose the <u>one</u> Axis II diagnosis most representative of the above case

7. Paranoid PD
8. Narcissistic PD
9. Schizoid PD
10. Avoidant PD
11. Schizotypal PD
12. Dependent PD
13. Antisocial PD
14. Obsessive-Compulsive PD

15. Borderline PD
16. Passive-Aggressive PD
17. Histrionic PD
18. Depressive PD

Rate how confident you are in your Axis I diagnosis for the Case.

1 2 3 4 5 6 7
Not at all Confident Very Confident

Rate how confident you are in your Axis II diagnosis for the Case.

1 2 3 4 5 6 7
Not at all Confident Very Confident

Rate the overall severity of this client.

1 2 3 4 5 6 7
Very Mild Very Severe

Rate what prognosis you would give this client.

1 2 3 4 5 6 7

Very Good Very Poor

Rate how likely this client would be to benefit from treatment.

1 2 3 4 5 6 7

Very Likely Very Unlikely

Rate the severity of each of the following symptoms using the following scale:

1 2 3 4 5 6 7

Absent Very Severe

___ recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide.
___ feelings of worthlessness or excessive or inappropriate guilt (which may be delusional) nearly every day (not merely self-reproach or guilt about being sick).
___ poor concentration or difficulty making decisions
___ low self-esteem
___ several discrete episodes of failure to resist aggressive impulses that result in serious assaultive acts or destruction of property.

___ the degree of aggressiveness expressed during episodes is grossly out of proportion to any precipitating psychosocial stressors.
___ marked distress that is in excess of what would be expected from exposure to the stressor.
___ significant impairment in social or occupational functioning.
___ inflated self-esteem or grandiosity
___ increase in goal-directed activity (either socially, at work or school, or sexually) or psychomotor agitation.
___ excessive involvement in pleasurable activities that have a high potential for painful consequences (e.g., the person engages in unrestrained buying sprees, sexual indiscretions, or foolish business investments).
___ is preoccupied with unjustified doubts about the loyalty or trustworthiness of friends or associates.
___ reads hidden demeaning or threatening meanings into benign remarks or events.
___ takes pleasure in few, if any, activities.
___ neither desires nor enjoys close relationships, including being part of a family.
___ behavior or appearance is odd, eccentric, or peculiar
___ odd beliefs or magical thinking that influences behavior and is inconsistent with subcultural norms (e.g., superstitiousness, belief in clairvoyance, telepathy, or "sixth sense").
___ failure to conform to social norms with respect to lawful behaviors as indicated by repeatedly performing acts that are grounds for arrest.
___ consistent irresponsibility, as indicated by repeated failure to sustain work behavior or honor financial obligations.
___ frantic efforts to avoid real or imagined abandonment. Note: Do not include suicidal or self-mutilating behavior.
___ a pattern of unstable and intense interpersonal relationships characterized by alternating between extremes of idealization and devaluation.
___ identity disturbance: markedly and persistently unstable self-image or sense of self.
___ impulsivity in at least two areas that are potentially self-damaging (e.g., spending, sex, substance abuse, reckless driving, binge eating). Note: Do not include suicidal or self-mutilating behavior.
___ recurrent suicidal behavior, gestures, or threats, or self-mutilating behavior.
___ affective instability due to a marked reactivity of mood (e.g., intense episodic dysphoria, irritability, or anxiety usually lasting a few hours and only rarely more than a few days.
___ chronic feelings of emptiness.
___ inappropriate, intense anger or difficulty controlling anger (e.g., frequent displays of temper, constant anger, recurrent physical fights).
___ transient, stress-related paranoid ideation or severe dissociative symptoms.
___ is uncomfortable in situations in which he or she is not the center of attention.
___ has a style of speech that is excessively impressionistic and lacking in detail.
___ is interpersonally exploitative (i.e., takes advantage of others to achieve his or her own ends).

___ lacks empathy: is unwilling to recognize or identify with the feelings and needs of others.
___ is unwilling to get involved with people unless certain of being liked.
___ is preoccupied with being criticized or rejected in social situations.
___ needs others to assume responsibility for most major areas of her life.
___ feels uncomfortable or helpless when alone because of exaggerated fears of being unable to care for herself.
___ is overconscientious, scrupulous, and inflexible about matters of morality, ethics, or values (not accounted for by religious identification).
___ shows rigidity and stubbornness.
___ usual mood is dominated by dejection, gloominess, cheerlessness, joylessness, unhappiness.
___ self-concept centers around beliefs of inadequacy, worthlessness, and low-self esteem.
___ complains of being misunderstood and unappreciated by others.
___ voices exaggerated and persistent complaints about personal misfortune.

Appendix E

Selected Items from Impact Message Inventory (Keisler, 1987)

This inventory contains words, phrases and statements which people use to describe how they are emotionally engaged or impacted when interacting with another person.

You are to respond to this Inventory by indicating how accurately each of the items describes your reactions to the particular person under consideration. Respond to each item in terms of how precisely it describes the feelings this person arouses in you, the behaviors you want directed toward her when she's around, and/or the descriptions of her that come to mind when you're with her. Indicate how each item describes your reactions on the following scale: (1 = Not at all; 2 = Somewhat; 3 = Moderately so; 4 = Very much so).

First, imagine you are in this person's presence, interacting with her. Focus on the immediate reactions you would be experiencing. Then read each of the items and fill in the number on the separate answer sheet which best describes how you would be feeling and/or would want to behave if you were, at this moment, in the person's presence. There are no right or wrong answers since different people react differently to the same person.

Please indicate which of the following best describes your reaction to the client using the following scale: 1 = Not at all; 2 = Somewhat; 3 = Moderately so; 4 = Very much so.

WHEN I AM WITH THIS PERSON SHE MAKES ME FEEL . . .

1. Bossed around (DOMINANT)
   1 2 3 4
2. Distant from her (HOSTILE)
   1 2 3 4
3. Like an intruder (HOSTILE)
   1 2 3 4
4. Forced to shoulder all the responsibility (HOSTILE)
   1 2 3 4
5. As if she's the class clown (DOMINANT)
   1 2 3 4
6. Uneasy (HOSTILE)
   1 2 3 4
7. Annoyed (HOSTILE)
   1 2 3 4
8. Taken charge of (DOMINANT)
   1 2 3 4

WHEN I AM WITH THIS PERSON SHE MAKES ME FEEL THAT . . .

9. I want to tell her to give someone else a chance to make a decision (DOMINANT)
   1 2 3 4
10. I'm going to intrude (HOSTILE)
   1 2 3 4
11. I want to stay away from her (HOSITLE)
   1 2 3 4
12. I should tell her she's often quite inconsiderate (DOMINANT)
   1 2 3 4
13. I want to get away from her (HOSTILE)
   1 2 3 4

WHEN I AM WITH THIS PERSON IT APPEARS TO ME THAT . . .

14. She wants to be the center of attention (DOMINANT)
   1 2 3 4
15. She doesn't want to get involved with me (HOSTILE)
   1 2 3 4
16. She wants me to put her on a pedestal (DOMINANT)
   1 2 3 4
17. She'd rather be alone (HOSITLE)
   1 2 3 4
18. She thinks its everyone for herself or himself (HOSITLE)
   1 2 3 4
19. She wants to be the charming one (DOMINANT)
   1 2 3 4
20. She thinks she's always in control of things (DOMINANT)
   1 2 3 4
21. She weighs situations in terms of what she can get out of them (DOMINANT)
   1 2 3 4
22. She'd rather be left alone (HOSTILE)
   1 2 3 4
23. She's carrying a grudge (HOSTILE)
   1 2 3 4
24. She thinks other people find her interesting, amusing, fascinating, and witty (DOMINANT)
   1 2 3 4

Note. Scale names are listed in bold next to the item. The actual format of the items will be modified for use in this study.

Appendix F

Interpersonal Style Descriptions (Keisler, 1987)

Hostile: A person who ridicules, belittles, or depreciates other people. Usually she displays a contemptuous "chip on the shoulder" attitude. This is evident by her frequent use of sarcastic or biting type of humor and her criticism and defiance of persons in authority. She often makes unfavorable or hostile remarks about her peers, and belittles or criticizes the success and strengths of others. On the other hand, she shows impatience or intolerance of others' mistakes and weaknesses. In dealing with people, she often shows anger or irritability and is known to tell people "off" when annoyed.

Dominant: A person who makes decisions like what to do or where to go when with another person. She volunteers advice and information when people have decisions to make and often talks her friends into doing what she would like. She generally takes charge of things when with people and presently directs the activities of several clubs and associations to which she belongs. She generally seizes opportunities to instruct or explain things to others which often leads her to dominate conversations, interrupting when others speak, and "talking others down." Moreover, she bosses friends and associates around and exploits or manipulates others for her own needs.

Appendix G

Demographic Questionnaire

---

1. Please indicate your gender
___ Male
___ Female

2. Please indicate your ethnicity
___ Caucasian
___ African American
___ Hispanic
___ Asian
___ Native American
___ Other

3. Please indicate your age: _____

4. Please indicate the type of degree that you hold.
___ Ph.D.
___ Psy.D.
___ Ed.D.
___ Other

5. Please indicate what field your degree is in.
___ Clinical Psychology
___ Counseling Psychology
___ School Psychology
___ Education
___ Other

6. Please indicate the year that you received your doctorate: _____

7. Please indicate the years of clinical experience you have had since completing your doctoral degree: ____

8. Please indicate your theoretical orientation.
___ Cognitive-Behavioral
___ Eclectic or Integrative
___ Humanistic
___ Psychodynamic
___ Other

9. Please indicate the primary setting in which you work.
___ Community Mental Health Center
___ University Medical School
___ VA Medical Center
___ General Medical Hospital
___ Private Psychiatric Facility
___ State Psychiatric Facility
___ Private Practice
___ Correctional Facility
___ Academic Department
___ Other

10. Please indicate the percentage of time that you work with the following age groups.
___ Children (0-12)
___ Adolescents (13-17)
___ Young adults (18-24)
___ Adults (25-44)
___ Middle age (45-64)
___ Older adults (65+)

11. Please indicate your level of experience with using the DSM-IV

1 2 3 4 5 6 7

None Very Experienced

12. Please estimate the percentage of time that you spend in direct patient contact.
____ %

Please include any comments that you have about your participation in this study:

Appendix H

Letter of Invitation

Diagnosing Borderline Personality Disorder: The effect of therapists' negative emotional reactions on diagnostic judgments.

I would like you to participate in my dissertation research on diagnostic decision-making. If you agree to participate, you will be asked to read a case vignette and rate it on a variety of dimensional measures. Please do NOT consult the DSM-IV when completing the survey, as I am primarily interested in your initial reactions based on your clinical judgment and clinical experience. You will also be asked to provide some demographic information and answer some questions regarding your professional background. I think you will find participation to be an interesting and informative experience. I am sure that your schedule is very busy, but I hope that you will take a few moments to participate. This study should only take 10-15 minutes to complete.

This study is entirely voluntary and you are under no obligation to participate in the study. If you agree to participate you will have the opportunity to be entered in a raffle in which you will have a chance to win one of three $50 gift certificates to Amazon.com. The submission of your responses on the web page constitutes your consent to participate.

This study can be completed by going online at my web page at: _____. To ensure that only those invited to participate in this study are included, you will be asked to enter the following verification code when you sign on: _____.

If you have any questions about the study, you may contact me by email at pykwmayo@isugw.indstate.edu or my faculty sponsor, Dr. June Sprock, at pyjune@isugw.indstate.edu. You may also reach us by phone through Indiana State University Department of Psychology at: (812) 237-2445. This project has been approved by the Institutional Review Board at Indiana State University. All reasonable precautions will be taken to preserve each participant's confidentiality. When the participants submit their responses, the data will be automatically entered into a computer database maintained by the researcher. This database will be password protected so that only the researcher will have access to the information. Backup copies of the database will be made periodically and stored in a locked office maintained by the researcher. If you have any questions about your rights as a research participant, please contact the IRB at (812) 237-8217 or irb@indstate.edu.

Thank you in advance for your time and effort. I greatly appreciate your willingness to share your clinical insights and expertise by participating in this research. Your participation is needed for the study to be a success.

Keith Mayo, M. S.
Doctoral Candidate
Indiana State University

Appendix I

Symptom Ratings for Both Cases

Table I.1

Case One: Mean Symptom Ratings for the Nine BPD Symptoms and Symptoms From Other Diagnoses With Means Greater Than 4.0

Symptom	Mean (SD)
BPD Symptoms	
"inappropriate, intense anger or difficulty controlling anger"	5.91 (1.14)*
"a pattern of unstable and intense interpersonal relationships characterized by alternating between extremes of idealization and devaluation"	4.81 (1.99)*
"affective instability due to a marked reactivity of mood"	4.67 (1.83)
"frantic efforts to avoid real or imagined abandonment"	2.79 (1.96)
"chronic feelings of emptiness"	2.64 (1.77)
"identity disturbance: markedly and persistently unstable self-image or sense of self"	2.46 (1.75)
"transient, stress-related paranoid ideation or severe dissociative symptoms"	2.24 (1.64)
"impulsivity in at least two areas that are potentially self-damaging"	2.10 (1.67)
"recurrent suicidal behavior, gestures, threats, or self-damaging behavior"	1.74 (1.54)

Table I.1 (*Continued*)

Symptom	Mean (SD)
Other Highly Rated Symptoms	
"complains of being misunderstood or unappreciated by others" (Passive-Aggressive (Negativistic) PD)	5.53 (1.36)
"lacks empathy: is unwilling to recognize or empathize with the feelings and needs of others" (Antisocial PD)	5.28 (1.64)
"the degree of aggressiveness expressed during episodes is grossly out of proportion to any precipitating psychosocial stressors" (Intermittent Explosive Disorder)	5.22 (1.56)
"shows rigidity and stubbornness" (Obsessive-Compulsive PD)	4.82 (1.65)
"significant impairment in social or occupational functioning" (multiple diagnoses)	4.54 (1.80)
"reads hidden demeaning or threatening meanings into benign remarks or events" (Paranoid PD)	4.26 (1.81)
"marked distress that is in excess of what would be expected from exposure to the stressor" (Adjustment Disorder)	4.23 (1.89)

* Symptom of BPD intended to be in the vignette

Table I.2

Case Two: Mean Symptom Ratings for the Nine BPD Symptoms and Symptoms from Other Diagnoses with Means Greater than 4.0.

Symptom	Mean (SD)
BPD Symptoms	
"frantic efforts to avoid real or imagined abandonment"	5.94 (1.18)*

Table I.2 (*Continued*)

"affective instability due to a marked reactivity of mood"	4.41 (1.83)*
"a pattern of unstable and intense interpersonal relationships characterized by alternating between extremes of idealization and devaluation"	4.01 (2.11)
"chronic feelings of emptiness"	3.64 (2.04)
"inappropriate, intense anger or difficulty controlling anger"	3.12 (1.76)
"identity disturbance: markedly and persistently unstable self-image or sense of self"	3.12 (1.93)
"transient, stress-related paranoid ideation or severe dissociative symptoms"	2.13 (1.53)
"impulsivity in at least two areas that are potentially self-damaging"	1.91 (1.57)
"recurrent suicidal behavior, gestures, threats, or self-damaging behavior"	1.85 (1.69)

Table I.2 (*Continued*)

Symptom	Mean (SD)
Other Highly Rated Symptoms	
"marked distress that is in excess of what would be expected from exposure to the stressor" (Adjustment Disorder)	5.40 (1.40)
"is preoccupied with unjustified doubts about the loyalty and trustworthiness of friends or associates" (Paranoid PD)	4.74 (1.89)
"complains of being misunderstood or unappreciated by others" (Passive-Aggressive Negativistic PD)	4.74 (1.74)
"significant impairment in social or occupational functioning" (multiple diagnoses)	4.64 (1.56)
"feels uncomfortable or helpless when alone because of exaggerated fears of being unable to care for herself"	4.27 (1.80)
"lacks empathy: is unwilling to recognize or empathize with the feelings and needs of others" (Antisocial PD)	4.10 (1.88)
"low self-esteem" (Depressive PD)	4.10 (1.98)

* Symptom of BPD intended to be in the vignette

Appendix J

IMI Ratings for Both Cases

Table J.1

Case One: Highest Rated IMI Items by Scale (M > 2.5) *and Average IMI Scale Scores*[1].

Dominant Scale		Hostile Scale	
Item	Mean (SD)	Item	Mean (SD)
"She wants to be the center of attention"	2.74 (1.03)	"She's carrying a grudge"	3.26 (.85)
		"Annoyed"	2.90 (.87)
"She weighs situations in terms of what she can get out of them"	2.70 (1.02)	"She thinks it's everyone for him or herself"	2.74 (1.01)
"She thinks she's always in control of things"	2.65 (1.00)	"I want to get away from her"	2.69 (.97)
		"I want to stay away from her"	2.66 (.93)
		"Distant from her"	2.61 (.93)
		"Uneasy"	2.54 (.88)
Average Scale Score	2.16 (.54)	Average Scale Score	2.30 (.49)

[1] All ratings were based on a scale from 1 (not at all so) to 4 (very much so).
Note. IMI items are paraphrased.

Table J.2

Case Two: Highest Rated IMI Items by Scale (M > 2.5) and Average IMI Scale Scores[1].

Dominant Scale		Hostile Scale	
Item	Mean (SD)	Item	Mean (SD)
"She wants to be the center of attention"	2.80 (.98)	"Forced to shoulder all the responsibility"	2.80 (1.01)
		"Annoyed"	2.51 (.86)
Average Scale Score	1.87 (.59)		1.95 (.48)

[1] All ratings were based on a scale from 1 (not at all so) to 4 (very much so).

Note. IMI items are paraphrased.

Appendix K

Relationship of IMI Items and BPD Symptoms to Participant Characteristics for Both Cases

Table K.1

Case One: Relationship of BPD Symptoms with Years of Clinical Experience and Percent of Direct Patient Contact

Years of Clinical Experience		
Symptom	*r*	*p*
"transient, stress-related paranoid ideation"	.23	.02
Percent of Direct Patient Contact		
Symptom	*r*	*p*
"affective instability"	.35	<.01
"identity disturbance"	.32	.01
"transient, stress-related paranoid ideation"	.26	.03

Note. BPD symptoms are paraphrased.

Table K.2

Case One: Relationship of IMI Items with Years of Clinical Experience and Percent of Direct Patient Contact

Years of Clinical Experience		
Item	*r*	*p*
"when I am with this person she makes me feel uneasy."	.24	.02
"when I am with this person it appears to me that she doesn't want to get involved with me"	.22	.03
"when I am with this person it appears to me that she'd rather be left alone"	.24	.02
Percent of Direct Patient Contact		
Symptom	*r*	*p*
"when I am with this person it appears to me that she doesn't want to get involved with me."	.24	.04
"when I am with this person it appears to me that she'd rather be alone"	.29	.02
"when I am with this person it appears to me that she thinks its everyone for himself or herself"	-.29	.02
"when I am with this person it appears to me that she thinks she's always in control of things"	.27	.03
"when I am with this person it appears to me that she'd rather be left alone"	.40	<.01

Table K.3

Case Two: Relationship of BPD Symptoms with Years of Clinical Experience and Percent of Direct Patient Contact

Years of Clinical Experience		
Symptom	*r*	*p*
"transient, stress-related paranoid ideation"	.28	.01
"identity disturbance"	.22	.03
Percent of Direct Patient Contact		
Symptom	*r*	*p*
"identity disturbance"	.24	.05
"transient, stress-related paranoid ideation"	.37	<.01

Note. BPD symptoms are paraphrased.

Table K.4

Case Two: Relationship of IMI Items and Percent of Direct Patient Contact[1]

Percent of Direct Patient Contact		
Symptom	*r*	*p*
"when I am with this person she makes me feel distant from her"	.26	.03
"when I am with this person she makes me feel like an intruder"	.28	.02
"when I am with this person she makes me feel taken charge of"	.26	.03

[1]No significant relationships were found between IMI items and years of clinical experience.

Appendix L

Order Effects

Table L.1

Case One: Effect of Diagnostic and Symptom Rating Order on BPD Symptom Ratings

	Order		
	Diagnosis, Symptom	Symptom, Diagnosis	
Symptom	Mean (SD)	Mean (SD)	F
abandonment	2.94 (2.07)	2.65 (1.86)	.53
unstable relationships[a]	4.28 (2.06)	5.29 (1.80)	6.79*
identity disturbance	2.68 (1.73)	2.25 (1.75)	1.46
impulsivity	2.51 (1.94)	1.73 (1.27)	5.70*
suicidal behavior, gestures	2.04 (1.92)	1.47 (1.01)	3.48
affective instability	4.60 (1.85)	4.75 (1.82)	.16
feelings of emptiness	2.74 (1.85)	2.55 (1.70)	.30
inappropriate intense anger[a]	5.77 (1.30)	6.04 (1.00)	1.41
paranoid ideation	2.68 (1.90)	1.84 (1.30)	6.78*

[a] Symptom of BPD intended to be in the Vignette

*$p < .05$

Note. BPD symptoms are paraphrased.

Table L.2

Case One: Effect of Case Order on BPD Symptom Ratings

	Case Order		
	Case 1, Case 2	Case 2, Case 1	
Symptom	Mean (SD)	Mean (SD)	F
abandonment	3.34 (2.15)	2.27 (1.63)	7.74**
unstable relationships[a]	5.02 (1.88)	4.61 (2.08)	1.06
identity disturbance	2.85 (2.00)	2.10 (1.40)	4.71*
impulsivity	2.43 (2.04)	1.80 (1.17)	3.50
suicidal behavior, gestures	2.15 (1.89)	1.37 (1.00)	6.62*
affective instability	4.57 (1.93)	4.76 (1.74)	.26
feelings of emptiness	2.87 (1.99)	2.43 (1.53)	1.53
inappropriate intense anger[a]	5.81 (1.15)	6.00 (1.13)	.69
paranoid ideation	2.62 (1.80)	1.90 (1.40)	4.85*

[a] Symptom of BPD intended to be in the vignette
*$p < .05$
**$p < .01$
Note. BPD symptoms are paraphrased.

Table L.3

Case Two: Effect of Case Order on BPD Symptom Ratings

	Case Order		
	Case 1, Case 2	Case 2, Case 1	
Symptom	Mean (SD)	Mean (SD)	F
abandonment[a]	6.04 (1.08)	5.84 (1.27)	.69
unstable relationships	3.17 (1.95)	4.80 (1.96)	16.86*
identity disturbance	3.09 (2.01)	3.16 (1.87)	.03
impulsivity	1.81 (1.58)	2.00 (1.56)	.36
suicidal behavior, gestures	2.06 (1.81)	1.65 (1.56)	1.50
affective instability[a]	3.87 (1.90)	4.92 (1.63)	8.57*
feelings of emptiness	3.70 (2.16)	3.59 (1.94)	.08
inappropriate intense anger	2.60 (1.70)	3.61 (1.67)	8.73*
paranoid ideation	2.04 (1.61)	2.22 (1.47)	.30

[a] Symptom of BPD intended to be in the vignette
*$p < .01$
Note. BPD symptoms are paraphrased.

Appendix M

Correlational and Logistic Regression Analyses for Case One

Table M.1

Correlations Between Criterion Variables and IMI Dominant and Hostile Scales.

	Dominant Scale	Hostile Scale
Criterion Variables		
BPD Assigned Diagnosis	-.12	.04
BPD Representativeness Ratings	.04	.23*
Average BPD Symptom Ratings	.23*	.35**
NPD Assigned Diagnosis	.08	-.02
NPD Representativeness Ratings	.44**	.36**

** $p < .01$
*$p < .05$
Note. Average BPD symptom rating was used as a predictor as well as a criterion variable in the regressions. NPD = Narcissistic Personality Disorder.

Table M.2

Predictive Value of Dominant and Hostile Scales for Diagnosis of BPD

Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	1.10	4.07	.04	.33
Hostile Scale	1.01	2.82	.09	2.75

Table M.3

Predictive Value of Dominant Scale, Hostile Scale, and Average BPD Symptom Ratings for Diagnosis of BPD

Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	-1.10	4.03	.05	.33
Hostile Scale	.84	1.80	.18	2.31
Average BPD Symptom Rating	.11	1.13	.29	1.11

Table M.4

Predictive Value of Dominant Scale, Hostile Scale, Average BPD Symptom Ratings, and Participant Gender for Diagnosis of BPD

Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	-1.19	4.23	.04	.31
Hostile Scale	.94	1.96	.16	2.56
Average BPD Symptom Rating	.11	1.08	.30	1.11
Participant Gender	1.22	7.75	.01	3.40

Appendix N

Multiple Regression Analyses for Case One

Table N.1

Predictive Value of Dominant and Hostile Scales for BPD Representativeness Ratings

Predictors	β	*t*	*p*
Dominant Scale	-.21	-1.59	.12
Hostile Scale	.37	2.78	.01

Table N.2

Predictive Value of Dominant Scale, Hostile Scale, and Average BPD Symptom Ratings for BPD Representativeness Ratings

Predictors	β	*t*	*p*
Dominant Scale	-.21	-1.63	.11
Hostile Scale	.24	1.84	.07
Average BPD symptom ratings	.36	3.58	< .01

Table N.3

Predictive Value of Dominant Scale, Hostile Scale, Average BPD Symptom Ratings, and Case Order for BPD Representativeness Ratings

Predictors	β	*t*	*p*
Dominant Scale	-.13	-.97	.33
Hostile Scale	.19	1.39	.17
Average BPD symptom ratings	.31	3.12	< .01
Case Order	-.19	-1.95	.05

Table N.4

Predictive Value of Dominant and Hostile Scales for BPD Symptom Ratings

Predictors	β	*t*	*p*
Dominant Scale	-.02	-.15	.88
Hostile Scale	.37	2.81	.01

Appendix O

Predictive Value of IMI Scales for Narcissistic Personality Disorder (NPD) Diagnosis and NPD Representativeness Ratings for Case One

Table O.1

Predictive Value of Dominant and Hostile Scales for Diagnosis of Narcissistic Personality Disorder (NPD).

Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	.71	1.43	.23	2.03
Hostile Scale	-.61	.85	.36	.54

Table O.2

Predictive Value of Dominant and Hostile Scales for Narcissistic Personality Disorder (NPD) Representativeness Ratings

Predictors	β	*t*	*p*
Dominant Scale	.36	2.87	.01
Hostile Scale	.11	.91	.37

Appendix P

Relationship Between IMI Scales and Ratings of Severity, Prognosis, and Likelihood of Treatment Benefit for Case One

Table P.1

Correlations Between IMI Scales and Ratings of Severity, Prognosis, and Likelihood of Treatment Benefit

	Dominant Scale	Hostile Scale
Criterion Variables	*r*	*r*
Severity	.30**	.26*
Prognosis	.19	.30**
Likelihood of Treatment Benefit	.19	.30**

Note. Higher values indicate worse prognosis and less likelihood of treatment benefit.
***$p<.001$
**$p < .01$
*$p < .05$

Table P.2

Dominant and Hostile Scales as Predictors of Severity, Prognosis, and Likelihood of Treatment Benefit

Severity: R = .309 (R^2 = .096, Adjusted R^2 = .076); F (2, 96) = 4.98, p = .01

Predictors	β	*t*	*p*
Dominant Scale	.23	1.71	.09
Hostile Scale	.10	.77	.44

Prognosis: R = .298 (R^2 = .089, Adjusted R^2 = .069); F (2, 96) = 4.57, p = .01

Predictors	β	*t*	*p*
Dominant Scale	-.03	-.19	.85
Hostile Scale	.31	2.32	.02

Likelihood of Treatment Benefit: R = .300 (R^2 = .090, Adjusted R^2 = .071); F (2, 96) = 4.71, p = .01

Predictors	β	*t*	*p*
Dominant Scale	-.02	-.17	.87
Hostile Scale	.32	2.36	.02

Note. Higher values indicate worse prognosis and less likelihood of treatment benefit.

Appendix Q

Predictors of Over- and Under-diagnostic Bias for Case One

Table Q.1

Predictive Value of Dominant and Hostile Scales for Over-diagnostic Bias

Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	-.92	2.72	.10	.40
Hostile Scale	.32	.60	.60	1.38

Table Q.2

Predictive Value of Dominant and Hostile Scales for Under-diagnostic Bias

Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	.30	.15	.70	1.35
Hostile Scale	1.05	1.46	.23	2.85

Table Q.3

Predictive Value of Dominant Scale, Hostile Scale, and Order of Diagnostic and Symptom Ratings for Diagnostic Bias

Over-Diagnostic Bias

Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	-.95	2.89	.09	.39
Hostile Scale	.25	.17	.68	1.28
Rating Order	1.02	5.15	.02	2.77

Under-diagnostic Bias

Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	.35	.16	.69	1.42
Hostile Scale	1.11	1.41	.24	3.03
Rating Order	-1.46	4.05	.04	.23

Table Q.4

Predictive Value of Dominant Scale, Hostile Scale, and Participant Gender on Diagnostic Bias

Over-Diagnostic Bias				
Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	-.97	2.83	.09	.38
Hostile Scale	.33	.28	.60	1.39
Clinician Gender	.92	4.27	.04	2.51
Under-Diagnostic Bias				
Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	.24	.09	.76	1.27
Hostile Scale	1.11	1.50	.22	3.04
Clinician Gender	-1.41	3.84	.05	.25

Appendix R

Correlational and Logistic Regression Analyses for Case Two

Table R.1

Correlations Between Criterion Variables and IMI Dominant and Hostile Scales.

Criterion Variables	Dominant Scale	Hostile Scale
BPD Diagnosis	.38**	.26*
BPD Representativeness Ratings	.45***	.35***
Average BPD Symptom Ratings	.41***	.46***
DPD Diagnosis	-.48***	-.26
DPD Representativeness Ratings	-.06	.07

Note. Average BPD symptom rating was used as a predictor as well as a criterion variable in the regressions. DPD = Dependent Personality Disorder.
***p<.001
**p < .01
*p < .05

Table R.2

Predictive Value of Dominant and Hostile Scales for Diagnosis of BPD

Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	1.40	7.34	.01	4.04
Hostile Scale	.15	.07	.80	1.17

Table R.3

Predictive Value of Dominant Scale, Hostile Scale, and Average BPD Symptom Rating for Diagnosis of BPD

Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	1.20	5.22	.02	3.31
Hostile Scale	-.35	.29	.59	.71
Average BPD Symptom Rating	.35	8.42	<.01	1.42

Table R.4

Predictive Value of Dominant Scale, Hostile Scale, and Case Order for Diagnosis of BPD

Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	1.10	4.08	.04	3.00
Hostile Scale	-.16	.00	.98	.98
Case Order	1.47	7.74	.01	4.33

Appendix S

Multiple Regression Analyses for Case Two

Table S.1

Predictive Value of Dominant and Hostile Scales for BPD Representativeness Ratings

Predictors	β	*t*	*p*
Dominant Scale	.39	3.29	<.01
Hostile Scale	.10	.88	.38

Table S.2

Predictive Value of Dominant Scale, Hostile Scale, and Average BPD Symptom Ratings for BPD Representativeness Ratings

Predictors	β	*t*	*p*
Dominant Scale	.30	2.74	.01
Hostile Scale	-.04	-.36	.72
Average BPD symptom ratings	.45	4.75	< .001

Table S.3

Predictive Value of Dominant Scale, Hostile Scale, Average BPD Symptom Ratings, and Case Order for BPD Representativeness Ratings

Predictors	β	*t*	*p*
Dominant Scale	.24	2.16	.03
Hostile Scale	-.06	-.57	.57
Average BPD symptom ratings	.45	4.87	<.001
Case Order	.19	2.15	.03

Table S.4

Predictive Value of Dominant and Hostile Scales for BPD Symptom Ratings

Predictors	β	*t*	*p*
Dominant Scale	.22	1.93	.06
Hostile Scale	.31	2.65	.01

Appendix T

Predictive Value of IMI Scales for Dependent Personality Disorder (DPD) Diagnosis and DPD Representativeness Ratings for Case Two

Table T.1

Predictive Value of Dominant and Hostile Scales for Diagnosis of Dependent Personality Disorder (DPD).

Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	-2.50	13.80	<.001	.08
Hostile Scale	.52	.62	.43	1.72

Table T.2

Predictive Value of Dominant and Hostile Scales for DPD Representativeness Ratings

Predictors	β	*t*	*p*
Dominant Scale	-.16	-1.25	.21
Hostile Scale	.17	1.34	.19

Appendix U

Relationship Between IMI Scales and Ratings of Severity, Prognosis, and Likelihood of Treatment Benefit for Case Two

Table U.1

Correlations Between IMI Scales and Ratings of Severity, Prognosis, and Likelihood of Treatment Benefit

	Dominant Scale	Hostile Scale
Criterion Variables	*r*	*r*
Severity	.37**	.35**
Prognosis	.32**	.31**
Likelihood of Treatment Benefit	.14	.29**

Note. Higher values indicate worse prognosis and less likelihood of treatment benefit.

**$p < .01$

Table U.2

Dominant and Hostile Scales as Predictors of Severity, Prognosis, and Likelihood of Treatment Benefit

Severity: R = .401 (R^2 = .161, Adjusted R^2 = .143); F (2, 96) = 9.10, p <.001

Predictors	β	*t*	*p*
Dominant	.25	2.09	.04
Hostile	.19	1.60	.11

Prognosis: R = .346 (R^2 = .119, Adjusted R^2 = .101); F (2, 96) = 6.44, p < .01

Predictors	β	*t*	*p*
Dominant	.20	1.62	.11
Hostile	.18	1.48	.14

Likelihood of Treatment Benefit: R = .291 (R^2 = .085, Adjusted R^2 = .065); F (2, 96) = 4.30, p = .02.

Predictors	β	*t*	*p*
Dominant	-.07	-.55	.58
Hostile	.38	2.61	.01

Note. Higher values indicate worse prognosis and less likelihood of treatment benefit.

Appendix V

Predictors of Over- and Under-diagnostic Bias for Case Two

Table V.1

Predictive Value of Dominant and Hostile Scales for Over-diagnostic Bias

Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	.91	2.40	.12	2.47
Hostile Scale	-.65	.74	.39	.52

Table V.2

Predictive Value of Dominant and Hostile Scales for Under-diagnostic Bias

Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	-.35	.31	.58	.71
Hostile Scale	1.08	2.10	.15	2.93

Table V.3

Predictive Value of Dominant and Hostile Scales and Rating Order for Diagnostic Bias

Over-Diagnostic Bias

Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	.91	2.41	.12	2.49
Hostile Scale	-.63	.68	.41	.53
Rating Order	-.13	.05	.83	.88

Under-Diagnostic Bias

Predictors	*B*	Wald	*p*	Exp(*B*)
Dominant Scale	-.35	.32	.57	.71
Hostile Scale	1.04	1.95	.16	2.84
Case Order	.21	.15	.70	1.24