Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Joanne Lisa Peart July 1, 2009

The Validity of Prototype Diagnosis in Everyday Practice

By

Joanne Lisa Peart
Doctor of Philosophy

Psychology

_____
Drew Westen, Ph.D.
Advisor

_____
Lawrence Barsalou, Ph.D.
Committee Member

_____
Edward Craighead, Ph.D.
Committee Member

_____
Scott Lilienfeld, Ph.D.
Committee Member

_____
Phillip Wolff, Ph.D.
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D. Dean of the Graduate School

_____
Date

The Validity of Prototype Diagnosis in Everyday Practice

By

Joanne Lisa Peart
M.A., Emory University, 2006
A.B., Harvard College, 2001

Advisor: Drew Westen, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the Graduate School of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Psychology
2009

Abstract

The Validity of Prototype Diagnosis in Everyday Practice
By Joanne Lisa Peart

The goal of the current study was to test the validity of a prototype-matching approach to clinical diagnosis in a naturalistic outpatient sample of 84 patients taken from 6 sites. Validity was determined by 1) correlating prototype ratings made by the treating clinician based on all available clinical data with patient self-reports of the same or similar constructs, and 2) assessing whether prototypes showed incremental validity above and beyond categorical *DSM-IV* diagnoses on a subset of disorders studied (two common mood disorders: major depressive disorder and dysthymic disorder). Significant correlations between prototypes and self-reports fell in the range of .22 to .48 for similar constructs. In a number of cases, associations between relevant self-report measures and diagnostic prototypes outperformed those associations between the same self-report measures and categorical diagnosis. Furthermore, a series of hierarchical linear regressions showed incremental validity of prototype diagnosis in a subset of mood disorders in predicting many clinically-relevant variables. Strong associations between prototypes and related self-report constructs coupled with generally equal or stronger associations when compared to *DSM-IV* categorical diagnosis and incremental validity in predicting criterion variables suggest that prototype diagnosis is a valid alternative to the categorical diagnostic approach that has been in place since *DSM-III*.

The Validity of Prototype Diagnosis in Everyday Practice

By

Joanne Lisa Peart
M.A., Emory University, 2006
A.B., Harvard College, 2001

Advisor: Drew Westen, Ph.D.

A dissertation submitted to the Faculty of the Graduate School of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Psychology
2009

Acknowledgements

℘ First, I would like to acknowledge my doctoral advisor, Dr. Drew Westen, for his guidance on this project. I greatly appreciate his enthusiasm, insight, expertise and feedback. I also thank Drew for his generosity in providing lab resources to aid in study completion.

℘ Thank you to Drs. Larry Barsalou, Ed Craighead, Scott Lilienfeld, and Phil Wolff for serving on my dissertation committee. I appreciate their thoughtful comments at the genesis of this project and their contribution as readers.

℘ I would like to acknowledge my parents, Viola and Michael, without whom none of my educational successes would have been possible. Thank you Mum and Dad for instilling in me the value of education, independence, and perseverance.

℘ I recognize the members of the Westen Lab who worked tirelessly on various administrative elements of this project, especially Kate Thomas who was extremely generous with her time and, overall, indispensable.

℘ I would like to thank Pavel Blagov for discussing various elements of this project with me and also for his overall support and friendship. When I need a special power calculator in the middle of the night, I know who to turn to.

℘ I am grateful for the support and encouragement of my friends and colleagues, whom I respect and greatly admire. I am especially thankful for the unwavering support of Erin Collier, Preeti Chauhan, Jamie Schumpf, Valerie Dixon, Megan Olden, Rachel Ammirati, Hanan Trotman, Susan Evans, Harry Segal, and the Lawrence Family.

℘ Finally, I would like to acknowledge the clinicians and patients who participated in the present study.

Table of Contents

List of Tables

The Validity of Prototype Diagnosis in Everyday Practice

The accurate categorization and diagnosis of psychopathology is crucial for effective clinical work and research alike. Over the various iterations of the *Diagnostic and Statistical Manual* (*DSM*), researchers have struggled with the proverbial joint-carving of a variety of psychiatric syndromes. However, there is growing concern that the *DSM* is cumbersome to use, distant from clinical diagnosis, and hence either not used or not used in reliable and valid ways in clinical practice. Under the current system, psychopathology is conceptualized categorically, despite scarce evidence of discrete breaks in the distribution of psychiatric symptomatology (see First, 2005) and mounting evidence in favor of dimensional diagnoses (see Widiger & Clark, 2000; Krueger, Watson, & Barlow, 2005). Furthermore, subthreshold pathology is difficult to capture using *DSM-IV* (APA, 1994), potentially causing the loss of valuable clinical information (Westen & Arkowitz-Westen, 1998).

The current system proves even more problematic when complicated diagnostic pictures arise. Clinicians are less accurate and less confident when diagnosing atypical patients versus prototypical ones (Russell, 1991; Clarkin, Widiger, Frances, Hurt, & Gilmore, 1983; Horowitz, Post, French, Wallis, & Siegelman, 1981; Horowitz, Wright, Lowenstein, & Parad, 1981). Similarly, in both research and practice, not otherwise specified (NOS) categories are often as or more commonly diagnosed as formally-defined diagnoses and lack both coherence and prototypicality (First, 2005). The current method of counting criteria and applying arbitrary cut-offs, as well as requiring complicated decision rules, proves time-consuming and unparsimonious in clinical practice. In fact, research suggests that most clinicians do not make decisions this way

and instead tend to diagnose in a gestalt manner (Lipkowitz & Idupuganti, 1985; Jampala, Sierles, & Taylor, 1988).

Despite a plethora of research on how best to classify disorders (e.g., which criteria to include or exclude to maximize the distinctiveness of major depressive disorder vs. dysthymic disorder), surprisingly little work has focused on how best to implement a diagnostic system once the diagnoses have been refined. In other words, most research has focused on how best to classify psychopathology but very little has focused on how to make a reliable and valid diagnostic system that is clinically useful and user-friendly. Is there a way to revise the current system that better balances the need for validity, reliability, and clinical utility?

The current study attempts to address the chasm between formal classification and clinical practice by testing the construct validity of a prototype-matching approach. In a prototype-matching approach as tested here, descriptive paragraphs are used to capture the essence of a disorder rather than lists of symptoms used to define discrete diagnostic categories. Clinicians then make dimensional diagnoses, rating the extent to which the patient's pathology matches each prototype. For purposes of description, parsimony, and communication, the upper range of ratings (e.g., 4-5 on a 5-point scale) can be considered categorical diagnoses, thus permitting both dimensional and categorical diagnosis within the same system.

I begin by briefly reviewing how well the current diagnostic system is working in practice and research. Next, I examine the relevant theories of classification and categorization as they might apply to an operationalizable and science-based system of diagnosis such as prototype-matching. Third, I describe how prototype diagnosis has

fared in the research literature to date. Finally, I present the current study along with potential limitations and implications.

The Current Diagnostic System

The seminal versions of the *DSM* (*DSM I* and *II;* APA 1952, 1968) contained short descriptions of psychiatric conditions as observed by mental health professionals. However, after growing criticism that both the diagnoses and the etiological theories often built into them were not empirically-based, reliable, or valid (e.g., Skinner, 1986), later versions of the *DSM* moved away from purely clinically-derived diagnostic categories (that often required inferences about underlying processes not shared by clinicians of alternative theoretical perspectives and clinical judgment about the presence or absence of a given disorder in a given patient) in favor of a system based on directly observable symptomatology (e.g., Schneiderian taxonomies). The *DSM* has since been through a number of iterations, each with the hopes of bringing us closer to a valid and reliable nosological system. While the manual has improved inarguably over the years, there is a new surge of dissatisfaction over its limitations. Tellingly, Verheul (2005) found in a systematic review of the current categorical system as well as alternative dimensional systems, that the categorical system "[had] the least evidence for clinical utility, especially with respect to coverage, reliability, subtlety, and clinical decision-making" (p. 295).

A principal concern with the current manual is that it treats mental disorders categorically, despite accumulating evidence in favor of dimensional diagnosis for most forms of psychopathology (e.g., Widiger & Clark, 2000; Livesley, Schroeder, Jackson, & Jang, 1994; also see Widiger & Samuel, 2005). So important is the question of whether

to conceptualize mental disorders as sets of dimensions rather than categories that it was identified as one of seven major nomenclature issues requiring attention for *DSM-V* (see Skodol & Bender, 2009).

A number of problems are inherent in categorical approaches. A primary difficulty is the high degree of diagnostic comorbidity. Many *DSM* disorders are comorbid (a term now widely used simply to refer to co-occurring disorders; see Lilienfeld, Waldman, & Israel, 1994), which suggests poor discriminant validity or common dimensions not well accounted for by discrete categories. In fact, more often than not, psychiatric disorders are comorbid. Co-occurrence occurs at a rate that exceeds the joint probability of each disorder co-occurring by chance (Widiger & Samuel, 2005), and rates further soar when considering *lifetime* comorbidity (Brown, Campbell, Lehman, Grishman, & Mancill, 2001). The comorbidity problem is particularly well documented in the personality disorder literature. Indeed, the comorbidity of virtually every Axis I disorder with Axis II disorders is upwards of 50% (Westen, Heim, Morrison, Patterson, & Campbell, 2002). At a recent *DSM-V* conference, members of a special task force reached a consensus that at least Axis II of the *DSM-V* should be organized by dimensions (see http://dsm5.org/conference13.cfm), and researchers are increasingly suggesting the same for most Axis I disorders, such as mood and anxiety disorders (Brown & Barlow, 2005**)** and even psychotic disorders (Tsuang, Stone, & Faraone, 2000).

The overlap between normal personality and both Axis I and II psychopathology (Clark, 2005) creates a particularly strong case in support of dimensionality. Clark (2005) posits that the overlap is best explained by temperament, which serves as a

diathesis. Similarly, researchers have argued that antisocial behavior and substance abuse (Krueger, Markon, Patrick, & Iacono, 2005) as well as mood and anxiety disorders (Watson, 2005) are best conceptualized by hierarchical, dimensional models of externalizing and internalizing pathology (Kendler et al., 2003), respectively. Furthermore, taxometric analyses suggest that continuous dimensions underlie worry (Ruscio, Borkovec, & Ruscio, 2001) and post-traumatic stress disorder (Ruscio, Ruscio, & Keane, 2002). Findings on the continuity of depression have been more variable; however, a number of studies suggest continuity (e.g., Solomon, Haaga, & Arnow, 2001; Haslam & Beck, 1994; Ruscio & Ruscio, 2002; Ruscio & Ruscio, 2000; Whisman & Pinto, 1997).

A second difficulty with the current diagnostic system is the presence of boundary disputes. Many researchers have bemoaned the seemingly arbitrary distinctions made between diagnostic categories that produce not only inflated estimates of comorbidity but gaps in the psychiatric nomenclature. Examples of disputed territories include bipolar II (which bridges the gap between bipolar I and cyclothymia) and mixed anxiety depressive disorder (which straddles the mood and anxiety disorders). Countless other disputed areas abound the *DSM*, stretching across the entire range of recognized psychopathology (see Widiger & Samuel, 2005). These "boundary disputes" have led to a proliferation of NOS diagnoses designed to capture symptomatology that does not meet *DSM* diagnostic criteria but is significant enough to warrant diagnosis. Widiger and Samuel (2005) call this a "wastebasket category," which highlights the hodgepodge of subthreshold and atypical cases that end up being "dumped" in the NOS bin. The NOS category is often used when existing *DSM* categories are not adequate (Clark, Watson, & Reynolds, 1995),

further suggesting that the current system has limited utility. This can also lead to mental illness going undiagnosed, unrecognized, and untreated, as has been the case for many patients with subthreshold disorders such as subthreshold depression (Pincus, McQueen, & Elinson, 2003). Subthreshold depression is a highly prevalent disorder (Cuijpers, Smit, & van Straten, 2007), associated with significant functional impairment, increased utilization of medical services (Wagner et al., 2000), an increased mortality rate (Cuijpers & Smit, 2002; Cuijpers & Schoevers, 2004), and an increased risk for developing major depressive disorder (Cuijpers & Smit, 2004; Fergusson, Horwood, Ridder, & Beautrais, 2005).

A final limitation is that the current diagnostic method does not take into consideration the cognitive processing parameters of human diagnosticians (i.e., clinicians). Despite the 30 years of effort to hone the diagnostic lists comprising *DSM* criterion sets, a lack of precision in defining some forms of psychopathology coupled with over-precision in defining others leads to problematic inter-rater reliability and a lack of stability in diagnostic categories. The question of how to operationalize dimensional diagnosis to maximize its utility in everyday practice (one of the main purposes of the diagnostic manual) is rarely addressed.

Because the *DSM* is an atheoretical manual, disorders are presented as checklists of symptoms with no description of causal mechanisms, etiology, or mental processes characteristic of a given disorder other than a small handful of overt symptoms. These checklists lack any cognitive coherence and hence are difficult for clinicians to remember and rate reliably. Using this approach, clinicians cannot draw upon clinical experience, prior knowledge, or even scientifically well-corroborated evidence about patterns of

covariation or causation in making diagnoses, and diagnoses are made based on a symptom count cut-off, often irrespective of *which* symptoms are endorsed. As a result, the current system requires that symptoms that may be most central to a disorder are often given equal weight to those that are less important. Of course, there are a limited number of exceptions to this - such as diagnostic categories that require *certain* symptoms for a diagnosis or monothetic categories (e.g., anorexia nervosa) that require the endorsement of *all* symptoms in order to meet diagnostic criteria (which itself is problematic, e.g., inclusion of amenorrhea as necessary to the diagnosis when many patients with clear cases of anorexia nervosa are not, or not yet, amenorrheic).

While overcommitment to theory may encourage rigidity and confirmation bias, arbitrary cut-offs are equally problematic. Little balance exists in the *DSM* between limiting arbitrary inferences, halo effects, confirmation biases, and similar heuristics, on the one hand, and harnessing clinical expertise and the inherently dimensional nature of most psychopathology, on the other. A prototype approach may allow clinicians to piece together an understanding of a patient's pathology that provides the best "goodness-of-fit" to the data. The latter principle is consistent with the cognitive theory of explanatory coherence (Thagard, 1989; 2000), in which the mind is simultaneously calculating the likelihood of multiple possible explanations or ways of classifying data based on the gestalt or pattern of data, ultimately leading to judgments of goodness-of-fit that take into account all of the available evidence (including the greater or lesser strength of certain kinds of evidence). Further, Ahn and colleagues have shown that classification is strongly influenced by patterns of covariation that suggest causal relations in lay people (Kim & Ahn, 2002a) and clinicians (Kim & Ahn, 2002b) alike. Indeed, research on

categorization in cognitive science suggests that causal inferences are central to categorization in virtually all aspects of human cognition (see Murphy & Medin, 1985), all of which have been deliberately removed from the diagnostic method used in the *DSM*.

Do clinicians actually make decisions as prescribed in *DSM-IV*, counting up symptoms, paying equal consideration to each symptom? The limited research available (e.g., Jampala et al., 1988; Morey & Ochoa, 1989; Lipkowitz & Idupuganti, 1985) suggests otherwise. Notably, one study (Morey & Ochoa, 1989) found a discrepancy between clinically-based diagnoses and *DSM* criteria-based diagnoses in a staggering 72% of cases. Instead of counting symptoms, it seems that clinicians get the "gist" of patient symptomatology (e.g., the patient is feeling down, has a poor appetite, is sleeping more than usual, and is contemplating suicide) and then diagnose accordingly (the patient has major depressive disorder). The exact symptom count appears to have significantly less influence than the gestalt presentation of the patient. Further, many researchers contend that the "natural" way for clinicians to diagnose is through analysis of graded typicality or comparison to prototypes (see Livesely, 1985; Cantor & Genero, 1986; Cantor, Smith, French, & Mezzich, 1980; Horowitz et al., 1981) which also allows clinicians to weight phenomena.

Of course, diagnosis should not slavishly follow the way humans naturally categorize, which may suffer from a number of flaws, biases, and heuristics. Clinicians may not use the *DSM* as intended because it is time-consuming even though doing so would lead to more valid diagnoses (as hypothesis I addresses in this dissertation). However, user-friendliness should be one important criterion in construction of a

diagnostic manual intended for use by clinicians (First et al., 2004), particularly when growing evidence suggests that requiring clinicians to wade through hundreds of pages of diagnostic criteria to make diagnoses that could be made much more simply (and perhaps with as much or greater reliability and validity than the current approach) is increasing, not decreasing, the gulf between science and practice, as clinicians are not using the manual as intended.

Classification and Categorization

Though an ideal diagnostic system would balance heuristics, user-friendliness, and diagnostic accuracy, some inherent difficulties exist in doing so. It is widely known that individuals often neglect base rates in probability judgment, violating Bayesian rules of statistical prediction (i.e., people are most likely to ignore base rates and instead base probability judgments on new evidence; Kahneman & Tversky, 1973). Finn (1982) once argued that base rates should be considered when establishing psychodiagnostic rules, lest poor treatment decisions be made as a result. However, as Widiger (1983) pointed out, utility adjustments are unnecessary to maximize patient benefit, as treatment decisions are not based solely on diagnosis. Further, it is unclear that gestalt prototype judgments about the extent to which patients match a given diagnostic prototype are more vulnerable to failure to consider base rates than the current system, in which clinicians are forced to make categorical judgments about individual criteria with varying base rates (which they then add up to make diagnoses based on arbitrary cut-points that have been developed without any knowledge of underlying base rates of true taxa if such exist for a given disorder). Indeed, it is equally likely that requiring clinicians to make dimensional

judgments of goodness-of-fit that do not require categorical diagnosis may minimize errors stemming from failure to consider base rates.

Clinical decision-making often involves cognitive heuristics, the simple rules or "short cuts" that govern human decision-making, which have both adaptive value and also the potential for mistaken judgments (Kahneman, 2003; Gigerenzer, 2007). The representativeness heuristic (Kahneman & Tversky, 1973) occurs when someone makes a judgment about a person, object, or event by comparing it to other people, objects, or events. In other words, an item is judged based on its similarity to the population from which it is selected. Two elements determine this subjective probability: (i) the degree to which the item is similar in essential characteristics to its parent population and (ii) the degree to which it reflects the salient features of the process by which it is generated (Kahneman & Tversky, 1972). While the representative heuristic can lead to errors such as base-rate neglect, it can also aid in clinical decision-making.

Heuristics are not merely "shortcuts." They are useful and adaptive mechanisms that allow humans to make decisions rationally and quickly (see Kahneman & Tversky, 1973; Hutchinson & Gigerenzer, 2005; Gigerenzer, 2008). Research suggests that clinicians frequently rely on the representativeness heuristic when making clinical decisions (Dawes, 1986; Garb, 1996). Such heuristics occur, in large part, due to the limited information-processing capacities of humans coupled with the fact that "our minds are not built (for whatever reason) to work by the rules of probability" (Gould, 1992, p.469).

In the case of mental illness, clinicians may compare a presenting patient to the *typical* person with a given disorder or the *prototypical* patient with a given disorder

(Garb, 1996). In the former case, the clinician uses something closer to a *stereotype* (i.e., an overgeneralized account) to arrive at a conclusion. A stereotype, therefore, might be thought of as a kind of prototype, albeit an imprecise one – with a typicality assumption that is based on the experience of the clinician rather than the universe of psychopathology or sound scientific description that might be built into a standardized diagnostic prototype used in a diagnostic manual. In the case of a prototype, the strength of the association is the similarity between the stimulus (the patient) and the category prototype (prototypical patient). Note that this is different from exemplar theory in which the strength of an association is based on the sum of similarities between a given stimulus (the patient) and all encoded exemplars (all encoded patients in a certain category).

Much of the cognitive literature suggests that exemplar theories often provide more accurate classification than do prototype views (see Medin & Smith, 1984; Ross & Makin, 1999). While exemplar-based classification models predict appropriate use of base-rate information and most prototype models imply an insensitivity to base-rate information, a prototype approach to diagnosis may take advantage of the characteristic advantages of exemplars but in a more scientifically valid manner, by creating standardized diagnostic prototypes that are essentially aggregated exemplars across clinicians and settings.

The central question here is how to make use of clinician expertise and normative cognitive processes (which are generally adaptive) while limiting the maladaptive use of heuristics such as representativeness (which can be closer to stereotypes or the clinician use of individual, idiosyncratic prototypes based on their own experiences). Making ratings based on unstandardized prototypes can prove problematic in that it potentially

relies on the opinion, experience, and exposure of an individual clinician to particular kinds of patients. By way of illustration, clinicians' stereotypes are vulnerable to race and gender biases (e.g., Ford & Widiger, 1989; Loring & Powell, 1988) and have also been found to differ- sometimes significantly- from one clinician to the next (McFall, et al., 1991; Livesely et al., 1987; Blashfield & Haymaker, 1988). Conclusions drawn about "typical" antisocial behavior, for example, may be different if the clinician works in a prison setting versus a small community psychotherapy practice. In turn, this could potentially encourage the misuse of the availability heuristic, which occurs when people judge the frequency of events by the ease with which they come to mind (Tversky & Kahneman, 1973). While use of availability heuristics may allow clinicians to take advantage of base rate information, the base rates accessed may be idiosyncratic. Indeed, a community practitioner may have a lower threshold for antisocial behavior based on limited exposure and, therefore, a skewed normative reference point. Finally, and importantly, studies suggest that prototypes used by clinicians differ from *DSM* criteria (Blashfield & Haymaker, 1988; Livesly et al., 1987; McFall et al., 1991), further contributing to a dilution of diagnostic constructs.

Because much of categorization depends on prototypes, it would be useful (and potentially much more scientifically sound) for the diagnostic manual to provide clinicians with standardized, empirically derived prototypes rather than the idiosyncratic prototypes derived from clinicians' idiosyncratic experiences or particular theories of psychopathology that clinicians appear to use despite the presence of a standardized diagnostic manual that has existed for nearly 30 years but proven too cumbersome for clinical utility. (The prototypes that I used in the current study are abstract, idealized

examples of a category that reflect the best available evidence at the time of construction of the diagnostic manual; see Sprock, 2003; Livesley & Jackson, 1986; Blashfield, 1985.) This approach does not eliminate clinical judgment or use of clinical experience in making diagnostic judgments, with all the attendant advantages and limitations (see Westen & Weinberger, 2004). Rather, it uses standardized, scientifically-derived prototypes that can anchor clinicians' experiences, so that as they become more familiar with a diagnostic dimension, they theoretically should develop richer, not more idiosyncratic, diagnostic prototypes.

In fact, precisely what led to prototype theories of categorization was the realization of the problems with classical notions of categorization based on defining features (e.g., Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Cantor & Mischel, 1979; Mervis & Rosch, 1981). Most categories are difficult to define precisely with necessary and sufficient criteria (and this problem is only magnified when considering dimensions, such as diagnostic dimensions). When defining the category "birds," for example, a number of shared properties come to mind. For example, "flying animal," "has feathers", and "lays eggs." However, not all birds have all of these defining features. A penguin is a bird but does not fly. Furthermore, not all animals with these features are birds. For example, a platypus lays eggs but is a mammal.

The classical defining features approach is inadequate for many forms of categorization, including the diagnosis of psychopathology (see Cantor & Genero, 1986). Like the platypus, psychopathology often falls within "fuzzy categories" consisting of members who share many features but not a set of necessary and sufficient ones. For example, suicidal ideation is commonly associated with depression, but not all depressed

individuals have thoughts of ending their life, and conversely, many other disorders (e.g., anxiety disorders, substance use disorders, psychotic disorders) can lead to suicidality.

A number of substantial changes have occurred in the relevant categorization literature over the last few decades (Kim & Ahn, 2002a). Originally, most categorization theories focused on classical rule-based approaches in which categories had necessary-and-sufficient defining features (e.g., Medin, 1989). In fact, the first two editions of the *DSM* (APA, 1958; APA, 1968) used this approach. This type of rule-based approach lost favor due to the challenge of finding appropriate defining features. Instead, prototype approaches slowly gained popularity (e.g., Rosch, 1978; Millon & Davis, 1996). In the prototype approach, diagnostic categories are represented as prototypes or averaged, abstract representations of patients with a given disorder (Kim & Ahn, 2002b). A prototype-like approach was adopted beginning with *DSM-III-R* (APA, 1987) which allowed for more diagnostic flexibility. Prototypical patients for many disorders were assumed to have a number of checklist features; however, only some of those features were *necessary* for a diagnosis. This model is problematic for some disorders, however, because it requires a mixed model, incorporating elements of a defining features model *as well as* a prototype model. Specifically, the current system confuses matters in that it 1) requires dichotomous (Yes/No) classification of criteria and diagnosis and 2) provides a list of features presumed to define the disorder rather than exemplifying its most defining features. Central to a prototype approach is the ability to make judgments on the degree to which a case resembles the prototype. However, the current system forces dichotomous classification (for both individual symptoms and the overall diagnosis). Furthermore, the polythetic decision rules of the *DSM* provide disorder *definitions*

operationalized as non-coherent lists of attributes rather than *exemplars* or aggregated prototypical cases.

Prototype Literature to Date

While seldom utilized, prototype approaches have been used in a few other areas of psychological research, including personal relationships and attachment patterns (Frei & Shaver, 2002; Hassebrauck & Aron, 2001; Onishi, Gjerde, & Block, 2001; Klohnen & John, 1998). Millon and Davis (1996) proposed a series of personality prototypes that corresponded with Axis II of the *DSM-IV*, but were designed as heuristic constructs rather than rigid diagnostic categories. Further, Westen, Shedler, and Bradley (2006) found that use of personality disorder prototypes not only reduced comorbidity, as compared to diagnosing through the *DSM-IV*, but also offered slightly better predictive validity in predicting a variety of important criterion variables such as adaptive functioning, treatment response, and etiology. Unpublished data from the same research group (Westen, Bradley, & Hilsenroth) shows similar findings for mood, anxiety, eating, and childhood behavioral disorders.

Use of a proposed prototype approach could potentially improve diagnostic accuracy and reliability. Research suggests that atypical patients are diagnosed with less accuracy than typical ones (see Genero & Cantor, 1987; Clarkin, Widiger, Frances, Hurt, & Gilmore, 1983; Horowitz, Post, French, Wallis, & Siegelman, 1981; Horowitz, Wright, Lowenstein, & Parad, 1981; Russell, 1991). Sprock (2003) found higher inter-rater reliability for prototypic cases compared to non-prototypic cases using a categorical system but no differences in inter-rater reliability for prototypic and non-prototypic cases

using dimensional ratings. This highlights the potential inadequacy of categorical ratings for non-prototypical cases.

Empirical studies of prototype diagnosis are important as they could bring us one step closer to eliminating some of the artifacts associated with categorical diagnosis, such as diagnostic instability and unreliability, excessive comorbidity, and arbitrary thresholds. A prototype approach could also potentially minimize some of the individual variability in categorization decisions as a number of variables are thought to impact how humans categorize such as available cognitive resources (Devine, 1989), mood (Forgas, 1995a, 1995b; Isen, Miedenthal, & Cantor, 1992), and comparative context (Haslam, Oakes, McGarty, Turner, & Onorato, 1995).

The use of dimensional prototypes provides information above and beyond that of the current system. Rather than simply knowing whether a given disorder is present or absent, a prototype rating may provide information about the severity/magnitude of the disorder and capture "subthreshold" cases. This is particularly helpful, given that severity is a significant predictor of course, chronicity, and comorbidity (Clark, Watson, & Reynolds, 1995). Furthermore, a dimensional approach reduces unreliability, in part because continuous scores are more stable over time than dichotomous ones (Widiger & Clark, 2000) and less sensitive to minor changes in symptomatology. Indeed, the disappearance of one symptom could easily change a diagnosis from present to absent, whereas the recognition of small changes is intrinsic to dimensional approaches. While structured clinical interviews, such as the Structured Clinical Interview for *DSM-IV* (SCID-IV; First et al., 1996), frequently yield kappas of .80 and above, field trials tell a less reliable story. *DSM-IV* field trial studies have demonstrated poor inter-rater

reliability, test-retest reliability, and/or diagnostic stability in a number of disorders, including mood disorders (Keller et al., 1995), autism spectrum disorders (Mahoney et al., 1998), substance use disorders (Cottler et al., 1997), and personality disorders (Jane et al., 2006). However, research on personality disorders has shown that prototype diagnosis (one form of dimensional diagnosis) leads to inter-clinician reliabilities above $r = .70$ (Westen, Bradley, & Hilsenroth, unpublished data).

If, as many have argued, a dimensional approach to psychopathology may eventually replace the current categorical system, the new system must be empirically sound (i.e., valid and reliable) as well as practically sound (i.e., useful to clinicians). It should encompass the spectrum of psychopathology while providing at least as valid (and ideally more valid) information above and beyond the current categorical approach (Clark, Watson, & Reynolds, 1995; Sprock & Blashfield, 1991). Finally, such a system should be "user friendly"—at least user-friendly enough that clinicians actually can use it under the constraints imposed by clinical practice, time constraints imposed by third-party payers, and so forth—and facilitate communication among mental health professionals.

Irrespective of support in favor of dimensional models, some studies suggest that clinicians find categories to be more familiar and easier to use (e.g., Widiger, 1993; Widiger & Sanderson, 1995) and are more confident in making categorical ratings (Sprock, 2003). However, recent research on Axis II from three separate research groups (one headed by a cognitive scientist whose central area of research is categorization) finds that clinicians overwhelmingly prefer prototype diagnosis (even with unfamiliar diagnoses) to both trait approaches and to categorical diagnoses, as well as to

dimensionalized approaches that simply involve symptom counts instead of arbitrary cutoffs (Spitzer et al., 2008; Westen, Shedler, & Bradley, 2006; Rottman, Ahn, Sanislow, & Kim, 2009). Our lab has produced similar data for Axis I disorders, and has shown prototype diagnosis to have high inter-rater reliability based on two clinicians listening to the same initial psychotherapy hours (Westen et al., 2006).

While a number of concerns have been raised about the usability and practicality of dimensional approaches (summarized in First, 2005), many of these concerns are addressed by prototype diagnosis, which is a dimensional system but can also be used with ranges to indicate caseness, as in other areas of diagnosis, as is the case with intelligence quotients (IQ) and mental retardation. IQ is measured dimensionally, but the cut-off point for mental retardation is 70 - a point at which significant cognitive impairment occurs. Some researchers (e.g., Sprock, 2003; Widiger, 2000; Oldham & Skodol, 2000) have argued in favor of a hybrid model- combining dimensional and categorical elements- as a way of transitioning to dimensional approaches. Our method creates both dimensional and categorical diagnoses. A rating of "3" corresponds to the concept of "features" or subthreshold pathology. Scores of "4" or "5" indicate "caseness" or a categorical diagnosis.

Though prototype matching is one method at the forefront of dimensional approaches to psychopathology categorization (and is currently being considered by both the World Health Organization for its clinical manual of mental disorders and by the Axis II Work Group for *DSM-V*), as noted above, it is not the only way to make a dimensional diagnosis. In the field of personality disorders, for example, trait models have been proposed as a dimensional alternative to the current categorical system. Although these

models have a large body of scientific evidence behind them, they have not performed well in studies of reliability and clinical utility (see Spitzer et al., 2008). For example, researchers have shown particular interest in the Five Factor Model of personality (FFM; Costa & McRae, 1992) for the upcoming *DSM-V* (see Samuel and Widiger, 2006). The FFM provides a personality description based on 30 facets of personality that represent lower-order traits comprising 5 superordinate personality factors that have emerged across many forms of data and cultures (neuroticism, extraversion, openness, agreeableness, and conscientiousness). However, a recent set of studies (Rottman, Ahn, Sanislow, & Kim, 2009) found that FFM descriptors were insufficient in capturing the nuance of *DSM-IV* personality disorders. In these studies, not only were clinicians unable to reliably back-translate prototypic FFM profiles into *DSM-IV* disorders, but they also judged the FFM as lacking in clinical utility – much like the conclusions drawn in a consumer preference study of dimensional diagnosis by Spitzer and colleagues (2008).

The Present Study

The primary goal of the current study was to assess the construct validity of a prototype- matching approach to psychiatric diagnosis using a range of disorders as currently defined by *DSM-IV*. The diagnostic manual requires the clinician to evaluate the presence or absence of a varying number of criteria to arrive at a diagnosis. The current study evaluated a method in which the treating clinician read a brief paragraph containing a prototype of a given disorder, presented as a description of the syndrome in its "pure" or "ideal" form. Clinicians then evaluated the extent to which their patient matched the diagnostic prototype on a 1-5 scale (with 1 meaning "no match" and 5 meaning a "strong match"; see Appendix B). In this study, we derived the prototypes

from the criterion sets in *DSM-IV*, simply weaving them together in a way that maximized their narrative coherence, so that we would not be simultaneously varying both the underlying construct and the method of operationalizing it dimensionally. The ultimate goal of this line of research is to develop richer, more scientifically grounded, psychologically descriptive paragraphs (prototype descriptions) that capture the essence of a disorder in a coherent manner.

Hypotheses

As described in the prior sections, prototype diagnosis captures both subthreshold and atypical psychopathology. Furthermore, it allows human diagnosticians (i.e., clinicians) to diagnose more naturally and flexibly, permitting clinicians to make graded judgments of typicality based on judgments of goodness-of-fit or explanatory coherence. While there is limited research on the use of prototype diagnosis, existing research suggests that it aids in predictive validity and reliability, and decreases artifactual comorbidity. For these reasons, my main hypothesis was that the clinician prototype ratings for some of the highest-frequency disorders in the population from which the sample was drawn and which provide a wide sampling of nonpsychotic disorders (major depressive disorder, dysthymic disorder, bipolar disorder, generalized anxiety disorder, panic disorder, post-traumatic stress disorder, anorexia nervosa, and bulimia nervosa) would show convergent validity with self-report data on similar constructs provided by the patient (see Table 1). More specifically, I predicted that the correlations between diagnostic prototypes and related self-report measures would fall in the small to moderate range where constructs were not identical and moderate to large (e.g., $r = .30$ to .50) where constructs were essentially identical, establishing convergent validity. At the same time, I predicted small to null correlations between prototypes rated by clinicians and

self-reports of unrelated diagnostic constructs, establishing discriminant validity. Where disorders are typically comorbid (e.g., most mood and anxiety disorders), I predicted small to moderate correlations, which should nonetheless be smaller than the correlations between clinician- and patient-reports of the same constructs.

Table 1 contains a complete list of the predictions. Plus signs (+) indicate predictions for positive correlations in the small to moderate range, double plus signs (++) indicate predictions of positive correlations of a moderate to large magnitude, empty cells indicate no specific predictions or correlations that are expected to hover near 0.0 (e.g., *Personality Assessment Inventory (PAI) Depression* and the panic disorder prototype), and negative signs (-) indicate correlations expected to be negative (e.g., *Positive and Negative Affect Schedule-Trait Version (PANAS-T)Positive Affect* and the major depressive and dysthymic disorder prototypes) for discriminant validity.

Predictions were made based on prior research on construct similarity and dissimilarity and patterns of comorbidity commonly found in the research literature. For example, for the major depressive disorder prototype, I predicted a strong correlation with self-report measures and subscales central to depressive syndromes – *PAI Depression and PANAS-T Negative Affect.* I further predicted that the major depressive disorder prototype would show small to moderate correlations with other self-report measures and subscales tapping negative affect, such as *PAI Anxiety,* and *PAI Anxiety Related Disorders* and a negative correlation with *PANAS-T Positive Affect* and Global Assessment of Functioning (GAF) Scale scores. For the generalized anxiety disorder prototype, I expected strong associations with subscales related to the varying cognitive, affective, and physiological aspects of pervasive anxiety and negative affect, hence

positive correlations with *PAI Anxiety* (including cognitive, affective, and physiological subscales)*, PAI Anxiety Related Disorders,* and *PANAS-T Negative Affect* and a negative correlation with *PANAS-T Positive Affect* and GAF Scale scores.

As a secondary (exploratory) hypothesis, I predicted that diagnostic prototypes would show incremental validity in predicting the same criterion variables above and beyond categorical *DSM-IV* diagnoses. More specifically, I predicted that for several dependent variables of clinical importance (composite adaptive functioning by self report, composite adaptive functioning by clinician report, *PAI Depression* scales and subscales, and GAF), prototype diagnoses for major depressive disorder, dysthymic disorder, and bipolar disorder (the categorical *DSM* diagnoses for which we presented clinicians with a copy of the pages of the *DSM-IV* listing the criteria for the disorder and asked them to make a *DSM-IV* categorical diagnosis) would predict incremental variance in dependent variables, moving beyond construct validity to differential validity vis-à-vis the current categorical diagnostic system.

Method

Participants

Two categories of participants comprised this research study: 1) Patients receiving treatment in six locations a) The Emory University Psychological Center, which is the training site for Ph.D. students in clinical psychology, b) The Emory University Outpatient Psychotherapy Training Program, which is the outpatient training site for psychiatric residents, c) the Emory University Psychopharmacology Clinic, d) the Emory University Counseling Center, e) the Cambridge Health Alliance, and f) Grady Hospital; and 2) The clinicians at each of these locations (Ph.D. students in Clinical Psychology,

Psychiatric Residents, or associated faculty or trainees). All clinicians in training were advanced graduate students or psychiatry residents who were supervised by a licensed psychologist or psychiatrist. Data were obtained from 84 patients and their clinicians, which provided adequate power (.81) to identify moderate correlations of the magnitude I expected given cross-observer validation of prototype diagnosis at $p<.05$.

The patient participants consisted of both men (n=34) and women (n=50) of ages ranging from 18-60 years, M=37.9 (SD=12.3). Patients represented a wide range of socioeconomic and racial groups, but were predominantly Caucasian (79.5 %) and middle class (42.2 %). Further, patient participants showed a wide range in levels of functioning and degree of psychopathology, as evidenced by GAF Scale scores ranging from 28 (serious impairment) to 90 (good functioning), M=62.8, SD=10.8. The clinician participants were from one of three mental health subfields: psychiatry (24.1 %), psychology (55.4 %), and social work (20.5 %). For a detailed description of patient and clinician demographics and characteristics, see Table 2.

Measures

Clinicians completed prototype measures and provided categorical DSM-IV diagnoses for the three most prevalent, well-defined mood disorders (excluding NOS categories; major depressive disorder, dysthymic disorder, and bipolar disorder). We selected self-report measures that would assess the same or similar constructs as assessed by clinician prototype ratings. (A set of additional measures was included in the larger study, including measures of affect dysregulation and mood and anxiety disorder symptoms; however, only the subset of measures described below were included in the present study). Measures are listed below in order of administration.

Assessment Instruments

Clinician Report Measures

The ***Clinical Data Form—Clinician Form*** (*CDF-C*) is a clinician-report questionnaire, developed over several years, which assesses a range of variables relevant to demographics, diagnosis, and etiology. The CDF-C was included in order to measure clinician-reported patient global functioning. Clinicians first provided basic demographic data on themselves, including discipline (psychiatry or psychology), employment site, and sex; as well as the patient's age, sex, race, education level, socioeconomic status, etc. Following basic demographic and diagnostic questions, clinicians rated the patient's adaptive functioning (including the GAF Scale score). The *CDF-C* also assesses aspects of the patient's developmental and family history of potential relevance to the etiology of personality pathology and Axis I symptomatology. It has been used in a variety of empirical studies within our research group (e.g., Westen & Shedler, 1999).

Prototype Ratings of DSM-IV Disorders were provided for a number of Axis I disorders (major depressive disorder, dysthymic disorder, bipolar disorder, panic disorder, generalized anxiety disorder, post-traumatic stress disorder, anorexia nervosa, and bulimia nervosa). These prototypes were converted directly from *DSM-IV*. These particular Axis I disorders were chosen because of their prevalence in outpatient samples. For each disorder, clinicians rated the patient using a 5-point rating scale where "1" indicates little or no match to the disorder prototype and "5" indicates a very good match to the prototype.

The ***DSM-IV Axis I Mood Disorder Checklist*** is a present/absent checklist of diagnostic criteria from *DSM-IV* for three mood disorders (major depressive disorder,

dysthymic disorder, and bipolar disorder) taken from the *DSM-IV*. The checklist was included in order to have a categorical diagnosis to serve as a comparative basis for prototype matching. I asked clinicians to use those ratings to make formal diagnoses. I included this for preliminary analyses allowing, for these disorders, the assessment of incremental validity of prototype diagnosis over the dichotomous present/absent diagnoses clinicians make using *DSM-IV*, using the diagnostic manual as it is intended to be used (with the clinician actively checking off each criterion and applying diagnostic algorithms to make categorical diagnoses).

Patient Report Measures

*The **Clinical Data Form-Patient Version (CDF-P)*** is a patient-report version of a clinician-report questionnaire that assesses a range of variables relevant to demographics, diagnosis, and etiology. The CDF-P was included in order to measure self-reported patient global functioning. Following basic demographic and diagnostic questions, patients rate aspects of their adaptive functioning and developmental history of potential relevance to etiology.

*The **Positive and Negative Affect Schedule-Trait Version*** (*PANAS-T*; Watson, Clark, & Tellegen, 1998) is a 20-item self-report measure consisting of 10 positive affects (interested, excited, strong, enthusiastic, proud, alert, inspired, determined, attentive, and active) and 10 negative affects (distressed, upset, guilty, scared, hostile, irritable, ashamed, nervous, jittery, and afraid). Participants are asked to rate items on a scale from 1 to 5, based on the strength of emotion where 1 = "very slightly or not at all," and 5 = "extremely." Initial studies in development of the *PANAS-T* showed that the scales are stable at appropriate levels over a 2-month time period, highly internally

consistent and largely uncorrelated. The *PANAS-T* was included in order to provide validity for mood and anxiety disorder prototypes, given the negative and positive associations between mood disorders with positive and negative affectivity, respectively, and between anxiety disorders and negative emotionality.

*The **Anxiety Sensitivity Index*** (*ASI*; Reiss, Peterson, Gursky, & McNally, 1986) measures the fear of anxiety-related sensations, such as those common in panic disorder. It is a 16-item self-report questionnaire in which participants rate their fear of anxiety symptoms on a likert-type scale from 0 (very little) to 4 (very much). The *ASI* has been demonstrated to have good internal consistency (coefficient alpha = .79-.90) and good test-retest reliability (r=.75).

*The **Personality Assessment Inventory*** (*PAI*; Morey 1991) is designed to assess a wide range of psychopathology and contains 344 items that are answered on a four choice scale *(false, slightly true, mainly true, and very true)*. Items load onto 22 non-overlapping full scales under four domains: validity (4 scales), clinical (11 scales), treatment consideration (5 scales), and interpersonal style (2 scales). For the present study analyses, I included a subset of *PAI* clinical scales and subscales (anxiety, anxiety-cognitive, anxiety-physiological, anxiety-affective, anxiety related disorders, traumatic stress, depression, depression-cognitive, depression-affective, depression-physiological, mania, and suicidal ideation). The *PAI* has strong psychometric properties and a median alpha value of .86 for clinical populations.

*The **Eating Disorder Diagnostic Scale*** (*EDDS*; Stice, Telch, & Rizvi, 2000) is a 22-item self-report scale for diagnosing anorexia nervosa, bulimia nervosa, and binge-eating disorder. The *EDDS* symptom composite demonstrates strong test-retest reliability

(.87), internal consistency (.89) and convergent validity with other self-report measures as well as expert ratings of eating disorders.

Procedure

Clinicians received information about the study in several sites located in Atlanta at Emory University and Grady Hospital or in Cambridge, Massachusetts at the Program for Psychotherapy at Cambridge Health Alliance. We held a group meeting with a member of the study staff at each site. At this session we provided clinicians with an overview of the study goals, procedures, and questionnaires. Patients received information about the study from a Psychological Center, Psychopharmacology Clinic, or Psychiatry representative (a trained research assistant or unit administrative assistant). Potential participants were handed an information sheet describing the study and were instructed to inform the representative if they were interested in participating. If patients were interested in participating, representatives further explained the study.

While reviewing the patient information sheet with potential participants, trained study representatives were instructed to provide their patients with study information, including study goals and compensation. Patients who were willing to participate signed the informed consent form in the presence of a study representative and were given an envelope with the questionnaires at a time of convenience, usually before or after an appointment. The patient returned the packet of measures directly to the receptionist at the clinic or by mail, which triggered study personnel to contact the patient's clinician to complete a set of clinician report measures.

Data Analysis

The primary goal of the current study was to assess the construct validity of a prototype-matching approach to diagnosis with selected *DSM* disorders. Do *DSM* diagnoses assessed continuously as clinician prototypes correspond with continuous self-report measures of related constructs? To test my primary hypothesis, that the prototype-matching approach would demonstrate convergent and discriminant validity with patient self-report data, I used correlational analysis to create a multitrait-multimethod matrix to determine the association between diagnostic prototypes for which we had data and relevant self-report measures (see Table 3). As outlined in Table 1, I expected moderate to strong correlations between the prototypes and a number of scale and composite scores from self-report data. Data were collected from both clinicians and patients to minimize rater-based variance and to use a true multitrait-multimethod matrix for validity.

To test whether prototypes performed as well as categorical diagnoses for the mood disorder diagnoses on which we had complete data (prototype diagnoses, *DSM-IV* diagnoses made by the clinician when given the criteria directly from the diagnostic manual, and criterion measures such as the *PAI* and *PANAS* scales), I included a secondary set of analyses. Time constraints limited the number of disorders I could assess; therefore I chose three categorical diagnoses from the same diagnostic grouping (mood disorders): major depressive disorder, dysthymic disorder, and bipolar disorder. To determine the performance of the current diagnostic system, I reported the magnitude of raw correlations for associations between categorical clinician diagnoses of the aforementioned mood disorders and the same self-report variables used to assess the prototypes of the same disorders (e.g., *PAI* mood variables, *PANAS-T Negative Affect*).

Next, to determine whether categorical or prototype diagnosis outperformed the other, I compared raw correlations between prototypes and criterion variables with those between categorical *DSM-IV* diagnosis and criterion variables. If the prototype associations were to prove lower on average than correlations between categorical diagnoses and criterion variables, that would suggest that prototype diagnoses may have less validity than the carefully designed polythetic diagnoses used in *DSM-IV*, at least for these three disorders.

I then conducted a secondary incremental validity analysis using hierarchical multiple regression (entering categorical diagnoses, coded 0/1, in Step 1, and prototype diagnoses in Step 2) to determine whether the prototype method provided any information above and beyond the current system in predicting seven criterion variables (composite adaptive functioning-patient report, composite adaptive functioning-clinician report, GAF scale score, *PAI Depression*, *PAI Depression-Cognitive*, *PAI Depression-Affective*, and *PAI Depression-Physiological)*. Composite adaptive functioning-patient report was comprised of a range of self-report variables taken from the CDF-P including quality of romantic relationships rated on a 5-point scale from very poor (1) to loving and stable (5), quality of friendships, number of close relationships, history of abusive relationships, history of suicide attempts, history of psychiatric hospitalization, history of arrest, occupational functioning, and job loss. Composite adaptive functioning-clinician report was comprised of a range of clinician-reported variables taken from the CDF including GAF scale score, quality of romantic relationships, quality of friendships, employment history, of suicide attempts, history of psychiatric hospitalization, history of arrest, history of abusive relationships, occupational functioning, and number of close

relationships. I created composite scores by standardizing all variables before taking the average, so that no variable would have greater weight than any other.

Although I intended to study three disorders in this way, because of lack of variance in the sample, bipolar disorder was not included in the analyses (reflecting both the fact that participants were outpatients and that many of the clinics screened out patients likely to require hospitalization in the future). Were we to account for incremental variance in the criterion variables, that would provide an especially strong case for incremental validity, given that the prototypes are simply dimensionalized versions of the categorical diagnoses.

Results

Patient Sample Characteristics

The patient sample (*n*=84) consisted of 50 (59.5 %) women and 34 (40.5 %) men with a mean age of 37.9 years (*SD*= 12.3). The majority of patients were Caucasian (*n*=66, 79.5 %), but African-American (*n*=6, 7.2 %), Hispanic (*n*=1, 1.2 %), Asian (*n*=4, 4.8%), and "Other" (*n*=6, 7.2%) ethnic groups were also represented. Most patients were middle class (*n*=35, 42.2 %), with 25.3 % (*n*=21) rated as working class, 20.5 % (*n*=17) rated as upper-middle class, 9.6 % (*n*=8) rated as poor, and 2.4% (*n*=2) rated as upper class. Mean Global of Assessment of Functioning (GAF) score was 62.8 (*SD*=10.8) and the mean number of sessions seen in psychotherapy was 24.4 sessions (*SD*=18.4), suggesting that clinicians knew the patients considerably well.

Clinician Sample Characteristics

Clinicians (*n*=84) were comprised of psychiatry (*n*=20, 24.1 %), psychology (*n*=46, 55.4 %), and social work (*n*=17, 20.5 %) residents, trainees, and post-doctoral

practitioners. Clinicians saw patients at multiple sites: Cambridge Health Alliance (*n*=33, 39.3 %), the Emory University Outpatient Psychotherapy Training Program (*n*=20, 23.8 %), the Emory Psychological Center (*n*=24, 28.6 %), Grady Hospital (*n*=2, 2.4 %), and the Emory Psychopharmacology Clinic and Counseling Centers (*n*=2, 2.2 %). Three practitioners (3.6 %) did not report their site.

Validity of Prototypes

Construct Validity

To test my hypothesis that clinician-rated diagnostic prototypes would show convergent validity with similar patient self-report data constructs, I created a multitrait-multimethod matrix to determine relevant associations. Table 3 shows Pearson correlations addressing the relationship between prototypes and a range of self-report measures/criterion variables, many of which were significant. As predicted, the prototype for major depressive disorder (MDD) showed significant positive correlations with measures and subscales tapping various aspects of depressed mood, negative affect, and anxiety, notably the *PAI Depression* scale (including cognitive, affective, and physiological subscales), the *PANAS-T Negative Affect* scale, the *PAI Anxiety* scale (including cognitive, affective, and physiological subscales), the *PAI Anxiety Related Disorders* scale, and the *Anxiety Sensitivity Index*. The MDD prototype was also positively associated with *PAI Suicidal Ideation*. In addition, as expected, the MDD prototype was negatively correlated with the GAF Scale score as well as positive affect as measured by the *PANAS-T Positive Affect* scale.

Much like the MDD prototype, the dysthymic disorder prototype showed significant positive associations with measures related to negative affect, including the

PAI Depression scale (including the cognitive, affective, and physiological subscales), the *PANAS-T Negative Affect* scale, the *PAI Suicidal Ideation* scale, and the *PAI Anxiety- physiological* subscale. Furthermore, as predicted, the dysthymic disorder prototype was negatively correlated with the GAF Scale score as well as positive affect as measured by the *PANAS-T Positive Affect* scale. As expected, correlations between the dysthymic disorder prototype and criterion variables were lower than those between the MDD prototype and criterion variables – offering discrimination between the similar but distinct diagnoses. Contrary to prediction, the dysthymia prototype was not significantly associated with *PAI Anxiety* (or the cognitive and affective subscales) or *PAI Anxiety Related Disorders*, though these correlations largely approached significance. The bipolar disorder prototype, the final of the mood disorder prototypes, showed a significant positive association with the *PAI Aggression* but was not significantly associated with *PAI Depression*, *PAI Mania*, or GAF Scale scores as predicted. A lack of variance in this diagnosis likely contributed to a lack of significantly associated self-report scales, which will be further visited in the discussion session.

As hypothesized, the generalized anxiety disorder (GAD) prototype was associated with many measures related to various domains of anxiety and negative affect, particularly the *PANAS-T Negative Affect* scale, the *PAI Anxiety* scale (including cognitive and affective subscales), and the *PAI Anxiety Related Disorders* scale. Associations between the GAD prototype and the *PAI Anxiety- physiological* subscale closely approached significance (p=.05). Contrary to prediction, the GAD prototype was not significantly associated with the *Anxiety Sensitivity Index* or *PANAS-T Positive Affect*. The panic disorder prototype partially mirrored the positive associations of the GAD

prototype, with expected significant correlations with the *Anxiety Sensitivity Index*, the *PANAS-T Negative Affect* scale, the *PAI Anxiety* scale (including cognitive, physiological, and affective subscales), and the *PAI Anxiety Related Disorders* subscale. The prototype for post traumatic stress disorder (PTSD) was significantly positively associated with the *PAI Traumatic Stress* subscale and negatively associated with GAF scale scores.

Finally, the diagnostic prototype for bulimia nervosa showed a significant positive correlation with the total symptom count of the Eating Disorder Diagnostic Scale (*EDDS*) as well as a significant negative correlation with the GAF Scale scores. The prototype for anorexia nervosa was not significantly associated with any hypothesized self-report or clinician measures. Once more, it appears that inadequate variance in the sample contributed to the lack of association, which will be addressed as a study limitation and consideration for future research.

In addition to convergent validity between criterion variables and diagnostic prototypes, for the data also supported discriminant validity. For example, the majority of prototypes did not correlate with *PAI Traumatic Stress* subscale nor variables related to externalizing disorders (*PAI Aggression, PAI Alcohol Problems, PAI Drug Problems, and PAI Antisocial*), and those that did (e.g., PTSD prototype and *PAI Traumatic Stress*, bipolar prototype and *PAI Aggression*) were conceptually related.

Differential and Incremental Validity

Table 4 summarizes the association between the three *DSM-IV* categorical diagnoses (major depressive disorder, dysthymic disorder, and bipolar disorder) and the same measures utilized in the primary analysis. To compare the magnitude of the

associations between diagnostic prototypes and relevant self-report measures with those between categorical diagnoses and the same self-report measures, I used Fisher's z. Table 5 summarizes the Fisher's z comparisons.

Although the difference between many of the associations was marginal, the diagnostic prototypes showed significantly stronger correlations on a number of measures when compared to their categorical counterparts. The most compelling example of these can be seen in the dysthymic prototype, where associations with self-report measures outperformed categorical associations for six different measures and approached significance for four. Thus, the prototypes appear to be either equally or more strongly associated with relevant criterion variables, such as negative affect, than do categorical diagnoses.

Finally, for the two mood disorders on which variance was substantial enough to warrant a test, major depressive disorder and dysthymic disorder, I conducted a series of hierarchical multiple regressions (see Appendix A) to detect any incremental validity for prototype matching above and beyond *DSM-IV* categorical diagnosis in predicting 7 clinically relevant variables (composite adaptive functioning-patient report, composite adaptive functioning-clinician report, GAF scale score, *PAI Depression*, *PAI Depression-Cognitive*, *PAI Depression-Affective*, and *PAI Depression-Physiological*). To avoid spurious results reflecting the number of regressions, I only considered patterns, and considered analyses predicting three variables for each disorder to be primary (*PAI Depression*, GAF, and composite adaptive functioning-patient report) and the remainder secondary, with the expectation that both prototypes would capture subthreshold pathology that would predict self-reported depression and show an impact on global

functioning on measure of global functioning (GAF) familiar to clinicians and a patient-reported measure of global functioning completely independent of clinician diagnoses. For each analysis, I first entered *DSM-IV* categorical diagnosis in Step 1 and then added prototype diagnosis of the same disorder in Step 2 to see if prototype diagnosis explained incremental variance. I then reversed the order to see if categorical diagnosis contained information not contained in the prototype diagnosis of the same construct. These are highly conservative procedures given that the two predictor variables in each equation are highly correlated because they consist of precisely the same criteria except that they are presented and rated differently.

Prototype diagnosis for major depressive disorder provided incremental validity above and beyond categorical diagnosis for a range of relevant criterion variables. Specifically, addition of the major depressive disorder prototype to the major depressive disorder categorical diagnosis explained significant incremental variance in *PAI Depression*, *PAI Depression-Affective subscale*, and GAF scale score. Furthermore, incremental variance explained approached significance (p=.06) for *PAI Depression-Cognitive* and *PAI Depression-Physiological* subscales. When reversing the analyses and entering diagnostic prototypes for major depression first, significant incremental validity was explained by major depressive disorder categorical diagnoses in predicting *PAI Depression*, *PAI Depression Cognitive*, *PAI Depression Affective*, and composite adaptive functioning-patient report, with a trend towards incremental variance explained in predicting composite adaptive functioning-clinician report (p=.06)

Prototype diagnosis for dysthymic disorder also provided incremental validity above and beyond categorical diagnosis for a range of relevant criterion variables.

Specifically, addition of the dysthymic disorder prototype to the dysthymic disorder categorical diagnosis explained significant incremental variance in *PAI Depression*, *PAI Depression-Cognitive, PAI Depression-Affective, PAI Depression-Physiological*, GAF scale score, and composite adaptive functioning-patient report. When reversing the analyses and entering diagnostic prototypes for dysthymic disorder first, significant incremental validity was not explained by categorical dysthymia diagnoses in predicting any of the seven criterion variables.

Discussion

The question of whether mental illnesses are best represented as discrete conditions or continua has been a longstanding issue, present throughout the history of the *DSM* (Widiger & Simonsen, 2005). The last several editions of the *DSM* have conceptualized psychopathology categorically, though it is becoming increasingly evident that this approach is inadequate for both clinical and empirical purposes, given that most psychopathology is dimensionally distributed in nature.

The current categorical system of diagnosis has a number of limitations in conceptualizing the universe of psychopathology. Problems with artifactual comorbidity, cumbersome usage, poor detection of subthreshold cases, short-term diagnostic instability, and arbitrary diagnostic thresholds highlight the limitations of the *DSM*. Alternatively, more evidence and support is building in favor of dimensional models of psychopathology that view psychopathology as having a continuous distribution and, therefore minimizes excessive (i.e., artifactual) co-occurrence, allows for better detection of subthreshold cases, and offers a clinical utility and ease-of-use not currently enjoyed by *DSM* users. This study aimed to test the construct validity of one of those dimensional

approaches, a prototype matching approach, which provided clinicians with descriptions of psychopathology weaved into a narrative form taken directly from current *DSM-IV* criteria.

The first aim was to determine whether the prototypes were valid diagnostic constructs. I correlated diagnostic prototypes for eight Axis I disorders completed by treating clinicians (major depressive disorder, dysthymic disorder, bipolar I disorder, generalized anxiety disorder, panic disorder, post-traumatic stress disorder, anorexia nervosa, and bulimia nervosa) with a range of self-report measures and subscales as well as the Global Assessment of Functioning (GAF) Scale score. Overall, prototype diagnosis showed convergent validity with a range of relevant criterion variables. Significant associations fell mostly in the moderate range, lending initial support to the construct validity of the prototypes.

Patterns of correlations between diagnostic prototypes and relevant criterion measures were also conceptually sensible rather than haphazard. For example, the MDD prototype correlated most strongly with criterion variables related to negative emotionality or negative affectivity (Eysenck, 1967; Gershuny & Sher, 1998; Watson & Clark, 1984), a construct used to describe a predisposition to the experience of negative affective states such as anxiety, sadness, and depression. Indeed, the MDD prototypes correlated with criterion variables related to depression (including cognitive, physiological, and affective elements of depression), and showed secondary correlations with variables measuring various aspects of anxiety (including cognitive, physiological, and affective elements of anxiety), anxiety sensitivity, and suicidal ideation among others – highlighting that dimensions such as Negative Affectivity (see Krueger et al., 1998)

likely underlie comorbidity which is also largely present in the *DSM* diagnostic categories from which we created the prototypes . Prototypes also unveiled significant associations with the GAF Scale score, suggesting that prototypes not only correlated with measures of related psychopathology, but also clinically-meaningful variables such as those related to adaptive functioning. In the case of three disorders, bipolar disorder, anorexia nervosa, and bulimia nervosa inadequate variance in the sample likely contributed to an attenuation of associations, as no participants were given prototype ratings of 5 and few were even given ratings of 4, so that the distribution ranged from "does not resemble this prototype" to "some clinical features." This will be further discussed in the limitations section.

The data also suggest that the diagnostic prototypes possessed discriminant validity, in that they did not correlate with constructs with which they should be dissimilar. For example, the majority of prototypes did not correlate with *PAI Traumatic Stress* subscales, and the prototype that did (post-traumatic stress disorder) was theoretically related and therefore expected to correlate.

The second aim of the study was to see whether a prototype matching approach might show incremental validity beyond the current *DSM-IV* system. Though issues of incremental validity were secondary to the main analysis, I found that comparing associations between categorical diagnosis and criterion variables with associations between prototype diagnosis and the same criterion variables, in the vast majority of cases, showed that the prototype diagnosis had either equal associations or stronger associations than categorical diagnosis. A compelling example of this was found in the dysthymia prototype, which significantly outperformed the categorical dysthymia

diagnosis, which has practical significance due to the tendency of the *DSM* to miss subthreshold depressive symptomatology. The stronger performance of the prototype is also particularly striking, as all clinicians were familiar with *DSM-IV* categorical diagnosis, but not prototype diagnosis. Though the dysthymia prototype showed convergent validity with related depressive symptomatology and , to a lesser extent, symptomatology tapping negative affect (such as *PANAS-T-Negative Affect* and *PAI Anxiety-Affective*), it is important to note that discriminant validity was also established. For example, examination of unrelated constructs, such as externalizing pathology as measured by *PAI Drug Problems, PAI Alcohol Problems,* and *PAI Aggression,* did not show significant associations.

A series of hierarchical linear regressions showed incremental validity of prototype diagnosis in predicting many external criterion variables. Addition of the major depressive disorder prototype to the major depressive disorder categorical diagnosis explained significant incremental variance in *PAI Depression*, *PAI Depression-Affective subscale*, and GAF scale score and a trend towards incremental variance for *PAI Depression-Cognitive* and *PAI Depression-Physiological* subscales.. Additional analyses determined that significant incremental validity was explained by major depressive disorder categorical diagnoses over major depressive disorder prototype diagnosis in predicting *PAI Depression*, *PAI Depression Cognitive*, *PAI Depression Affective*, and composite adaptive functioning-patient report, with a trend towards incremental variance explained in predicting composite adaptive functioning-clinician report. These results suggest that, for major depressive disorder, prototype-matching and categorical diagnosis

both appear to pick up overlapping regions of variance on criterion variables but provide at least partially non-overlapping information on some of those variables.

However, for dysthymic disorder, prototype-matching demonstrated clear superiority as compared to *DSM* categorical diagnosis. Indeed, prototype diagnosis for dysthymic disorder provided incremental validity above and beyond categorical diagnosis for *PAI Depression*, *PAI Depression-Cognitive, PAI Depression-Affective, PAI Depression-Physiological*, GAF scale score, and composite adaptive functioning-patient report. Categorical dysthymia diagnosis, in contrast, did not show incremental validity on any criterion variable, suggesting that it is losing information relative to dimensional diagnosis.

These data suggest that, at least for the Axis I disorders I studied, prototype matching is a valid method of diagnosing psychopathology, with as strong or stronger correlations with criterion variables as the categorical approach used for the several years. Such a system provided incremental predictive validity above and beyond DSM categorical diagnosis, particularly for dysthymic disorder. This system also utilizes the flexibility of dimensional diagnosis, correlates in predictable way with related criterion variables, and performs as well or better than the current diagnostic system. This is particularly striking given the greater familiarity of the *DSM* to clinician participants and the fact that prototype diagnosis has been seen as a particularly useful approach for Axis II (personality pathology) that is likely to be built into *DSM-V* (Skodol & Bender, 2009). Further, it is notable that clinicians in this study were relatively inexperienced and hence had not only just been steeped in the *DSM* in their training (favoring the *DSM*-based approach) but did not have rich knowledge structures to "fill out" the prototypes in their

minds from having seen many patients with the various forms of psychopathology they were describing, which likely attenuated correlations between predictor and criterion variables, particular unfamiliar diagnostic prototypes. One of the advantages of giving clinicians standardized diagnostic prototypes in a diagnostic manual is in fact its ability to allow them to "hang" their experience on shared prototypes rather than to develop their own idiosyncratic ones from which they make diagnoses. The findings are also striking in that they represent cross-informant correlations, not the usual correlations between self-reported psychopathology (when patients report their symptoms in a structured interview to a trained interviewer) and self-reported psychopathology (by other self-reports), which does not allow researchers to tease apart true variance from method variance (informant-related variance).

Limitations

Through this research, I aimed to offer an initial test of an alternative to the current diagnostic system for Axis I disorders by assessing the construct validity of a prototype-matching approach to clinical diagnosis. While this study offered a promising new way to look at psychiatric diagnosis, there are several limitations to consider while interpreting the findings. First, because all patients were outpatients, there was some clear restriction of range, particularly on disorders such as bipolar disorder and anorexia nervosa. This, in turn, likely attenuated the resultant correlations. The disorders that lacked variance (such as anorexia nervosa or bipolar disorder) might have been better represented in an inpatient setting.

A second limitation is the relatively small sample size from which the data were gathered. Recruiting limitations did not allow for a larger sample, beyond the 84

clinicians and their patients. However, a power analysis determined that the sample had adequate power (.81) to identify moderate correlations of the magnitude I expected given cross-observer validation of prototype diagnosis at $p<.05$.

A third limitation involves the secondary analysis of incremental validity. As categorical data were only available on three mood disorders (major depressive disorder, dysthymic disorder, and bipolar disorder), limited conclusions could be made from subsequent analyses. Furthermore, because of insufficient variance in the bipolar diagnosis, the diagnosis was dropped from the exploratory analysis, further limiting conclusions that could be drawn regarding incremental validity. However, incremental validity was demonstrated in a number of the incremental validity analyses, showing that overall prototype diagnosis provides equal or greater explanatory variance as *DSM* categorical. While the analyses demonstrated incremental validity, future studies or replications should be attuned to the need for greater variance and, possibly, additional power. Despite these less than ideal parameters, it is important to note that prototype diagnosis performed as well as or better than the well-studied and highly familiar *DSM* criteria. This coupled with research suggesting that prototypes have more predictive validity than *DSM* personality disorder diagnoses (e.g., Westen, Shedler, & Bradley, 2006, 2004) and are the dimensional method found most clinically useful by practicing clinicians (Spitzer et al., 2008) suggests that prototype diagnosis is a worthy option to pursue.

A fourth limitation is that the criterion data are all self-reports, with the exception of the GAF scores. Although biases are inherent in all forms of data, it would have been useful to have such additional data sources as a structured interview (e.g., Structured

Clinical Interview for DSM Disorders; SCID) conducted by a second interviewer to provide a second set of criterion variables and, thus, data triangulation. (In unpublished data recently collected and analyzed from our lab using an inner-city African-American sample, we have found strong correlations between Axis I prototype diagnoses made using a systematic clinical interview and SCID diagnosis, further bolstering the findings here.)

A fifth potential criticism focuses on the comparison between continuous variables versus dichotomous variables. Can we be certain that the prototypes did not perform the same as any other dimensional approach would have, especially given that dimensional approaches tend to provide statistically more information? First, it is important to note that research has clearly demonstrated that not all dimensional approaches are equal, with prototype approaches substantially outperforming other dimensional approaches on measures of clinical utility for personality variables without sacrificing validity (Westen, Shedler, & Bradley, 2006; Rottman et al., 2009; Spitzer et al., 2008). In any case, my goal was not to demonstrate that prototype diagnosis is the only other possible dimensional approach to diagnosis, but that it could at least match categorical *DSM* diagnosis. It does, however, have multiple advantages from a clinical standpoint that have been empirically documented in multiple studies of clinical utility, including the fact that it is a dimensional system that still allows clinicians or researchers to communicate categorically when useful, as when they describe a patient as having "significant features" of dysthymic disorder, signified by a prototype rating of "3."

The final limitation regards the study clinicians, as many of the clinicians are in training or relatively inexperienced. Although clinicians who are earlier in their career

may not be as seasoned, in many respects this renders the data more conservative, as we would expect more experienced clinicians who have seen the disorders under study with greater frequency to be better able to use and recognize the prototypes, particularly in comparison to what are essentially the symptom checklists provided by *DSM-IV*.

Implications

The implications of an empirically valid prototype-matching approach to diagnosis are far-reaching. The current study addresses a growing interest in and belief in the potential of dimensional models of classification and diagnosis. As Rounsaville and colleagues (2002) stated, "There is a clear need for dimensional models to be developed and their utility compared with that of existing typologies in one or more limited fields, such as personality. If a dimensional system performs well and is acceptable to clinicians, it might be appropriate to explore dimensional approaches in other domains" (p. 13). The current study extends the existing dimensional diagnosis literature beyond the burgeoning personality disorders field and adds evidence of construct validity within the clinical disorders (Axis I). Ultimately, this adds to the previous dearth of systematic testing of operationalized dimensional diagnostic systems for Axis I.

A major advantage to the adoption of a prototype approach is that dimensional approaches can minimize artifactual comorbidity, the use of NOS categories, and the failure to capture subthreshold psychopathology. Both epidemiologic and clinical studies show staggering rates of comorbidity, both within and across Axes I and II that dilute the information that can be gleaned from multi-axial diagnosis. Furthermore, atypical or subthreshold cases are often banished to NOS categories or not captured, respectively. Prototype diagnosis allows much more flexibility in capturing disorders of varying

severity and prototypicality, thus preventing the loss of important clinical information and honoring the inherent heterogeneity of many *DSM* disorders. This clinical information can then be used for important pursuits, such as improving treatment specificity. Furthermore, a prototype approach would allow clinicians flexibility in diagnosis while avoiding the current quagmires of (sometimes) arbitrary differential diagnoses and honoring complicated decision rules.

Next is the potential to have a system in place that is balanced in terms of validity and utility (i.e., ease of use), thus narrowing the chasm between researchers and practitioners. In addressing diagnostic validity, various elements must be considered including face validity, descriptive validity, predictive validity, and, as was specifically addressed in the current study, construct validity (Blacker & Endicott, 2000). Though validity has often been discussed as an entirely separate domain from clinical utility, some have argued that they are concepts with considerable overlap (First et al., 2004), and Spitzer (2001) made a case that diagnostic validity should be understood as "the extent [to which] the defining features of a disorder provide useful information not contained in the definition of the disorder. (pp. 351)." Regardless of the degree of overlap, however, clinical utility should not an afterthought when it comes to revising the manual to be used by clinicians in diagnosing their patients. Indeed, from the seminal article by Robins and Guze (1970) to more recent appeals (see First et al., 2004) and even an introductory statement in the *DSM-IV* (APA, 1994), the importance of clinical utility for our diagnostic system has been heralded. This is especially important, given the consistently poor performance of the categorical system in comparison studies of clinical utility (Verheul, 2005; Spitzer et al., 2008).

Another important consideration is whether "consumers," be they clinicians or researchers, will find the proposed system to be helpful and user-friendly. Dimensional models are unfamiliar to those trained in a largely categorical system, which raises a question as to how practitioners would adapt to an unfamiliar system. Though concern over reaction to change in the current system is to be expected, it is not necessarily warranted or supported by the literature. Instead, studies suggest that clinicians find the prototype matching approach to be easier to implement and more clinically-meaningful than the *DSM* categorical system (Westen, Shedler, & Bradley, 2006) and even strongly preferred by clinicians asked to apply it to a specific case relative to alternative dimensional systems (see Spitzer et al., 2008). This finding has now been replicated by three different research teams, most recently by a team of cognitive scientists who are at the forefront of cognitive research on categorization (Rottman et al., 2009).

Because the prototype system differs from the current diagnostic system, many have naturally raised the question of how such a system would be put into place. Prototype diagnosis, especially the one proposed in the current study, is of high feasibility and could be implemented relatively easily with minor changes to the diagnostic manual. Although we would strongly advocate continued refinement of criteria, even weaving the current diagnostic criteria into coherent paragraphs and providing users with the 1-5 rating system ("no match" to "very good match"), as in the current study, would allow for a relatively easy transition. Westen and colleagues (2002) have also argued that prototype diagnosis could have even stronger clinical utility if patients receiving a prototype diagnosis of 2 or higher (i.e., those with some clinical features of a given disorder) were then rated on a series of further questions, such as

frequency, duration, age at which clinically significant features of the disorder emerged, and dimensions identified through factor analysis similarly rated on a simple 1-5 scale (e.g., the extent to which a depressed patient had cognitive symptoms, physiological symptoms, and melancholic features). Having clinicians only rate these additional dimensions on diagnoses applicable to the patient (e.g., who received a prototype rating of 2 or 3) would be an efficient way to collect additional data not currently provided by *DSM* diagnosis by obtaining highly relevant clinical information without overburdening clinicians by requiring that they make these additional ratings on diagnoses that do not apply to the patient.

Future Directions and Conclusions

The current study is by no means a definitive statement on dimensional diagnosis or prototype diagnosis but rather an effort to highlight the validity and potential utility of such an approach. As consideration is given to a dimensional transition in the future, there are a number of things to consider moving forward. One consideration is to determine whether or not prototype diagnosis is equally useful in the research domain. The extant literature has mostly evaluated the approach for clinical utility which is somewhat limited, as an ideal diagnostic system would have both clinical and research utility (although the *International Classification of Diseases* diagnostic system has produced parallel systems for clinicians and researchers, in recognition of their overlapping but disparate tasks). Second, there is a need for further study of a dimensional approach to Axis I disorders, as a preponderance of the already scant literature on dimensional diagnosis has been dedicated to personality disorders.

Future research should also test prototype matching on a larger and more varied scale. Use of inpatient sites as well as non-training sites could add to variability in terms of diagnosis, socioeconomic and racial background, as well as clinician characteristics. Addition of structured clinical interviews (e.g., SCID) could also provide data triangulation to further strengthen the validity of the approach.

Finally, additional research is needed on the validation of empirically-derived prototypes such as those proposed by Westen and colleagues (2006). In this case, empirical prototypes are not limited by preexisting *DSM-IV* criteria but are instead empirically-derived and may, therefore, have better convergent validity than prototypes woven out of diagnostic criteria that were products of committee consensus rather than statistical procedures such as factor analysis. Another potential consideration is the creation of diagnostic prototypes that include diagnostic exemplars. For example, a prototype for major depressive disorder might include a diagnostic prototype as proposed in the present study coupled with several exemplars (concrete representations, typical members). This could take advantage of the utility, stability, and potential incremental validity of prototypes while allowing diagnosticians to utilize typical members of the category that are likely to be committed to memory (exemplars). This suggestion is supported by research in the cognitive science literature showing that exemplars often outperform prototypes (see Nosofky and Zaki, 2002). Supplementing prototypes with exemplars would, in a single page, provide a highly useful training manual as well as a reliable, valid, and cognitively optimized method for diagnosis.

The timeliness of this research is particularly significant given the impending fifth edition of the *Diagnostic and Statistical Manual* and the *International Classification of*

Diseases, both of which are now considering prototype diagnosis at least for personality (and in the case of ICD, for the clinician version of the diagnostic system for all disorders). As task forces aim to formulate a system that allows for a valid and useful representation of the universe of psychopathology, a move away from tradition and towards innovation and growth is needed. The current research suggests that a move towards a prototype-matching approach may not only be at least as valid as the current categorical approach but also clinically-meaningful, with the potential to minimize the problems inherent in the current system.

References

American Psychiatric Association (1952). *Diagnostic and statistical manual of mental disorders.* Washington, DC: Author.

American Psychiatric Association (1968). *Diagnostic and statistical manual of mental disorders (2nd ed.).* Washington, DC: Author.

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

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

Blacker, D., & Endicott, J. (2000). Psychometric properties: concepts of reliability and validity, in Handbook of Psychiatric Measures. Washington DC, American Psychiatric Association, pp7-14.

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

Blashfield, R.K., & Haymaker,D. (1988). A prototype analysis of the diagnostic criteria for DSM-III-R personality disorders. *Journal of Personality Disorders, 2,* 272-280.

Brown, T. A., & Barlow, D. H. (2005). Dimensional versus categorical classification of mental disorders in fifth edition of the diagnostic and statistical manual of mental disorders and beyond: Comment on the special section [Special Issue]. *Journal of Abnormal Psychology, 114, 551-556.*

Brown, T. A., Campbell, L. A., Lehman, C. L., Grisham, J. R., & Mancill, R. B. (2001). Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large clinical sample. *Journal of Abnormal Psychology, 110,* 585–599.

Cantor, N., & Genero, N. (1986). Psychiatric diagnosis and natural categorization: A close analogy. In T. Millon & G. Klerman (Eds.), *Contemporary directions in psychopathology* (pp.233-256). New York: Guilford Press.

Cantor, N., & Mischel, W. (1977). Traits as prototypes: Effects on recognition memory. *Journal of Personality and Social Psychology, 35,* 38-48.

Cantor, N., Smith, E.E., French, R., & Mezzich, J. (1980). Psychiatric diagnosis as prototype categorization. *Journal of Abnormal Psychology, 89,* 181-193.

Clark, L.A. (2005). Temperament as a unifying basis for personality and psychopathology. *Journal of Abnormal Psychology, 114,* 505-521.

Clark, L.A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: Challenges to the current system and future directions. *Annual Review of Psychology, 46,* 121-153.

Clarkin, J.F., Widiger, T.A., Frances, A., Hurt, S.W., & Gilmore, M. (1983). Prototypic typology and the borderline personality disorder. *Journal of Abnormal Psychology, 92,* 263-275.

Costa, P, & McRae, R (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five Factor Inventory (NEO_FFI) Professional Manual. Odessa (FL): Psychological Assessment Resources.

Cottler, L.B., Grant, B.F., Blaine, J., Mavreas, V., Pull, C., Hasin, D., Compton, W.M., Rubio-Stipec, M., & Mager, D. (1997). Concordance of DSM-IV alcohol and drug use disorder criteria and diagnoses as measured by AUDADIS-ADR, CIDI and SCAN. Drug and Alcohol Dependence, 47, 195-205.

Cuijpers, P., & Schoevers, R.A. (2004). Increased mortality in depressive disorders: a review. *Current Psychiatry Report, 6,* 430-437.

Cuijpers, P., & Smit, F. (2004). Subthreshold depression as a risk indicator for major depressive disorder: a systematic review of prospective studies. *Acta Psychiatrica Scandinavica, 109,* 325-331.

Cuijpers, P., & Smit, F. (2002). Excess mortality in depression: a meta-analysis of community studies. *Journal of Affective Disorders, 72,* 227-236.

Cuijpers, P., Smit, F., & van Straten, A. (2007). Psychological treatments of subthreshold depression: a meta-analytic review. *Acta Psychiatrica Scandinavica, 115,* 434-441.

Dawes, R.M. (1986). Representativeness thinking in clinical judgment. *Clinical Psychology Review, 6,* 425-441.

Devine, P.G. (1989). Stereotype and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5-18.

Eysenck, H. (1967). *The biological basis of personality.* Springfield, Illinois: Charles C. Thomas.

Fergusson, D.M., Horwood, L.J., Ridder, E.M., & Beautrais, A.L. (2005). Subthreshold depression in adolescence and mental health outcomes in adulthood. *Archives of General Psychiatry, 62,* 66-72.

Finn, S. (1982). Base rates, utilities, and DSM-III: Shortcomings of fixed-rule systems of diagnosis. *Journal of Abnormal Psychology, 91,* 294-302.

First, M.B. (2005). Clinical utility: A prerequisite for the adoption of a dimensional approach in *DSM. Journal of Abnormal Psychology, 114(4),* 560-564.

First, M.B., Pincus, H.A., Levine, J.B., Williams, J., Ustun, B., & Peele, R. (2004). Clinical utility as a criterion for revising psychiatric diagnosis. *American Journal of Psychiatry, 161(6),* 946-954.

First, M.B., **Spitzer,** R.L., **Gibbon,** M., & **Williams,** J.B. (**1996**) *Structured clinical interview for DSM-IV axis I disorders*. American Psychiatric Publishing, Washington, DC.

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

Forgas, J.P. (1995a). Mood and judgment: the Affect Infusion Model (AIM). *Psychological Bulletin, 117,* 39-66.

Forgas, J.P. (1995b). Strange couples: Mood effects on judgments and memory about prototypical and atypical relationships. *Personality and Social Psychology Bulletin, 21,* 747-765.

Frei, J.R., & Shaver, P.R. (2002). Respect in close relationships: Prototype definition, self-report assessment, and initial correlates. *Personal Relationships, 9,* 121-139.

Garb, H. (1996). The representativeness and past-behavior heuristics in clinical judgment. *Professional Psychology: Research and Practice, 27(3),* 272-277.

Genero, N., & Cantor, N. (1987). Exemplar prototypes and clinical diagnosis: Toward a cognitive economy. Journal of Social & Clinical Psychology, 5, 59-78.

Gershuny, B.S., & Sher, K.J. (1998). The relation between personality and anxiety: Findings from a 3-year prospective study. *Journal of Abnormal Psychology, 107,* 252-262.

Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. New York: Viking Penguin.

Gigerenzer, G. (2008). Why heuristics work. *Perspectives on Psychological Science, 3(1),* 20-29.

Gould, S.J. (1992). *Bully for brontosaurus: Further reflections in natural history.* New York: Penguin Books.

Haslam, N., & Beck, A.T. (1994). Subtyping major depression: A taxometric analysis. *Journal of Abnormal Psychology, 103,* 686-692.

Haslam, S.A., Oakes, P.J., McGarty, C., Turner, J.C., & Onorato, R.S. (1995). Contextual changes in the prototypicality of extreme and moderate outgroup members. *European Journal of Social Psychology, 25,* 509-530.

Hassebrauck, M., & Aron, A. (2001). Prototype matching in close relationships. *Personality and Social Psychology Bulletin, 27(9),* 1111-1122.

Horowitz, L.M., Post, D.L., French, R. de S., Wallis, K.D., & Siegelman, E.Y. (1981). The prototype as a construct in abnormal psychology: 2. Clarifying disagreement in psychiatric judgments. *Journal of Abnormal Psychology, 90,* 575-585.

Horowitz, L.M., Wright, J.C., Lowenstein, E., & Parad, H.W. (1981). The Prototype as a construct in abnormal psychology: 1. A method for deriving prototypes. *Journal of Abnormal Psychology, 90(6),* 568-574.

Hutchinson, J.M.C., & Gigerenzer, G. (2005). Simple heuristics and rules of thumb: Where psychologists and behavioural biologists might meet. *Behavioural Processes*, *69(2),* 97-124.

Isen, A.M., Niedenthal, P.M., & Cantor, N. (1992). An influence of positive affect on social categorization. *Motivation and Emotion, 16,* 65-78.

Jampala, V., Sierles, F., & Taylor, M. (1988). The use of *DSM-III* in the United States: A case of not going by the book. *Comprehensive Psychiatry, 29*, 39-47.

Jane, J., Pagan, J., Turkheimer, E., Fiedlerd, E., Oltmanns, T. (2006). The interrater reliability of the Structured Interview for DSM-IV Personality. *Comprehensive Psychiatry, 47*, 368 – 375.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58 (9),* 697-720.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3,* 430-454.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 337–251.

Kendler, K.S., Prescott, C.A., Myers, J., & Neale, M.C. (2003). The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry, 60,* 929-937.

Keller, M., Klein, D., Hirschfeld, R., Kocsis, J., McCullough, J., Miller, I., First, M., Holzer 3rd, C., Keitner, G., & Marin, D. (1995). Results of the DSM-IV mood disorders field trial. *American Journal of Psychiatry, 152,* 843-849.

Kim, N.S., & Ahn, W. (2002a). The influence of naïve causal theories on lay diagnosis of mental illness. *American Journal of Psychology, 115,* 33-65.

Kim, N.S., & Ahn, W. (2002b). Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. *Journal of Experimental Psychology:General, 131(4),* 451-476.

Klohnen,E. & John, O.P. (1998). Working models of attachment: A theory-based approach. In J.A. Simpson and W.S. Rholes (Eds.), *Attachment theory and close relationships* (pp.115-140). New York:Guilford.

Krueger, R.F., Caspi, A., Moffitt, T.E., & Silva, P.A. (1998). The structure and stability of common mental disorders (DSM-III-R): a longitudinal-epidemiological study. *Journal of Abnormal Psychology, 107*, 216–227.

Krueger, R.F., Markon, K.E., Patrick, C.P., & Iacono, W.G. (2005). Externalizing psychopathology in adulthood: A dimensional spectrum conceptualization and its implications for *DSM-V*. *Journal of Abnormal Psychology, 114,* 537-550.

Krueger, R., Watson, D., & Barlow, D. (2005). Introduction to the special section: Toward a dimensionally based taxonomy of psychopathology. *Journal of Abnormal Psychology, 114(4),* 491-193.

Lilienfeld, S.O., Waldman, I.D., & Israel, A.C. (1994). Critical examination of the use of the term and concept of comorbidity in psychopathology research. *Clinical Psychology:Science and Practice, 1,* 71-83.

Lipkowitz, M.H., & Idupuganti, S. (1985). Diagnosing schizophrenia in 1982: The effect of DSM-III. *American Journal of Psychiatry, 142,* 634-637.

Livesley, W.J. (1985). The classification of personality disorder:I. The choice of category concept. *Canadian Journal of Psychiatry, 30,* 353-358.

Livesely, W.J., & Jackson, D.N. (1986). The internal consistency and factorial structure of behaviors judged to be associated with DSM-III personality disorders. *American Journal of Psychiatry, 143(11),* 1473-1474.

Livesley, W., Schroeder, M., Jackson, D., & Jang, K. (1994). Categorical distinctions in the study of personality disorder: implication for classification. *Journal of Abnormal Psychology, 103*, 6–17.

Livesley, W.J., Reiffer, L.T., Sheldon, A.E.R., & West, M. (1987). Prototypicality ratings of DSM-III criteria for personality disorders. *Journal of Nervous and Mental Disease, 175,* 395-401

Loring, M., & Powell, B. (1988). Gender, race, and the DSM-III: A study of the objectivity of psychiatric diagnostic behavior. *Journal of Health and Social Behavior, 29,* 1-22.

Mahoney, W.J., Szatmari, P., MacLean, J.E., et al. (1998). Reliability and accuracy of differentiating pervasive developmental disorder subtypes. *Journal of the American Academy of Child and Adolescent Psychiatry, 37*, 278–285.

McFall, M.E., Murburg, M.M., Smith, D.E., & Jensen, C.F. (1991). An analysis of criteria used by VA clinicians to diagnose combat-related PTSD. *Journal of Traumatic Stress, 4,* 123-136.

Medin, D. L. (1989). Concepts and conceptual structure. *American Psychologist*, *12*, 1469–1481.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology,* 35, 113–138.

Mervis, C.B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32,* 89-115.

Millon, T., & Davis, R.D. (1996). An evolutionary theory of personality disorders. In J.F. Clarkin and M.F. Lenzenweger (Eds.), *Major Theories of Personality Disorder* (pp.221-346). New York: Guilford Press.

Morey, L.C. (1991). *The Personality Assessment Inventory: Professional manual.* Odessa: FL: Psychological Assessment Resources.

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,* 180-192.

Murphy, G.L., & Medin, D.L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92,* 289-316.

Nosofsky, R. & Zaki, S. (2002). Exemplar and Prototype Models Revisited: Response Strategies, Selective Attention, and Stimulus Generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28(5),* 924-940.

Onishi, M., Gjerde, P.F., & Block, J. (2001). Personality implications of romantic attachment patterns in young adults: A multi-method, multi-informant study. *Personality and Social Psychology Bulletin, 27(9),* 1097-1110.

Oldham, J.M., & Skodol, A.E. (2000). Charting the future of Axis II. *Journal of Personality Disorders, 14,* 17-29.

Pincus, H.A., McQueen, L.E., & Elinson, L. (2003). Subthreshold mental disorders: Nosological and research recommendations. In K.A. Phillips, M.B. First, & H.A. Pincus (Eds.). *Advancing DSM: Dilemmas in psychiatric diagnosis* (pp. 129-144). Washington, DC: American Psychiatric Association.

Reiss, S., Peterson, R.A., Gursky, D.M., & McNally, R.J. (1986). Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. *Behavior Research and Therapy, 24,* 1-8.

Robins, E., & Guze, S. (1970). The establishment of diagnostic validity in psychiatric illness: its application to schizophrenia. *American Journal of Psychiatry, 126,* 983-987.

Rosch, E. (1978) Principles of categorization. In E. Rosch & B. Lloyd (Eds.), Cognition and categorization (pp. 27-47). Hillsdale, NJ: Erlbaum.

Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology,8,* 382-439.

Ross, B.H., & Makin, V.S. (1999). Prototype versus exemplar models. In R.J. Sternberg (Ed.) The nature of cognition (pp. 205-241). Cambridge, MA: MIT Books.

Rottman, B., Ahn, W., Sanislow, C., & Kim, N. (2009). Can clinicians recognize DSM-IV personality disorders from five-factor model descriptions of patient cases? *American Journal of Psychiatry,166,* 1-7.

Rounsaville, B., Alarcon, R., Andrews, G., Jackson, J., Kendell, R., & Kendler, K. (2002). Basic nomenclature issues for DSM-V, in a research agenda for DSM V. Edited by Kupfer, D., First, M., Regier, D. Washington DC, American Psychiatric Association. pp 1-29.

Ruscio, A.M., Borkovec, T.D., & Ruscio, J. (2001). A taxometric investigation of the latent structure of worry. *Journal of Abnormal Psychology, 110,* 413-422.

Ruscio, A.M., & Ruscio, J. (2002). The latent structure of analogue depression: Should the Beck Depression Inventory be used to classify groups? *Psychological Assessment, 14,* 135-145.

Ruscio, J., & Ruscio, A.M. (2000). Informing the continuity controversy: A taxometric analysis of depression. *Journal of Abnormal Psychology, 109,* 473-487.

Ruscio, A.M., Ruscio, J., & Keane, T.M. (2002). The latent structure of posttraumatic stress disorder: A taxometric investigation of reactions to extreme stress. *Journal of Abnormal Psychology, 111,* 290-301.

Russell, J.A. (1991). In defense of a prototype approach to emotion concepts. *Journal of Personality and Social Psychology, 60,* 37-47.

Skinner, H. A. (1986). Construct validation approach to psychiatric classification. In T. Millon & G. L. Klerman (Eds.), *Contemporary approaches to psychopathology* (pp. 307–330). New York: Guilford Press.

Skodol, A., & Bender, D. (2009). The Future of Personality Disorders in DSM-V? *American Journal of Psychiatry, 166(4),* 388-391.

Solomon, A., Haaga, D.A.F., & Arnow, B. A. (2001). Is clinical depression distinct from subthreshold depressive symptoms? A review of the continuity issue in depression research. *Journal of Nervous & Mental Disease*, *189*, 498-506.

Spitzer, R.L. (2001). Values and assumptions in the development of DSM-III and DSM-III-R: An insider's perspective and a belated response to Sadler, Hulhus, and Agich's "On values in recent American classification." *Journal of Nervous and Mental Disease, 189,* 351-359.

Spitzer, R.L., First, M.B., Shedler, J., Westen, D., & Skodol, A.E. (2008). Clinical Utility of 5-Dimensional Systems for Personality Diagnosis: A "Consumer Preference" Study. *Journal of Nervous and Mental Disease, 196(5),* 356-374.

Sprock, J. (2003). Dimensional versus categorical classification of prototypic and nonprototypic cases personality disorder. *Journal of Clinical Psychology, 59(9),* 992-1014.

Sprock, J., & Blashfield, R.K. (1991). Classification and nosology. In M. Hersen, A.E. Kazdin, & A.S. Bellack (Eds.), The clinical psychology handbook (2nd ed., pp. 329–344). New York: Pergamon Press.

Stice, E., Telch, C. F., & Rizvi, S. L. (2000). Development and validation of the Eating Disorder Diagnostic Scale: A brief self-report measure of anorexia, bulimia, and binge-eating disorder. *Psychological Assessment, 12,* 123–131.

Samuel, D., & 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.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences, 12,* 435-467.

Thagard, P. (2000). *Coherence in thought and action.* Cambridge, MA: MIT Press.

Tsuang, M.T., Stone, W.S., & Faraone, S.V. (2000). Toward reformulating the diagnosis of schizophrenia. *American Journal of Psychiatry, 157(7)*, 1041-1050.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5,* 207-232.

Verheul, R. (2005). Clinical utility for dimensional models of personality pathology. *Journal of Personality Disorders, 19,* 283-302.

Wagner, H.R., Burns, B.J., Broadhead, W.E., Yarnall, K.S.H., Sigmon, A., Gaynes, B.N. (2000). Minor depression in family practice: functional morbidity, comorbidity, service utilization and outcomes. *Psychological Medicine, 30,* 1377-1390.

Watson, D. (2005). Rethinking the mood and anxiety disorders: A quantitative hierarchical model for *DSM-V*. *Journal of Abnormal Psychology, 114,* 522-536.

Watson, D. & Clark, L.A. (1984). Negative Affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin, 96,* 465-490.

Watson, D., Clark, L.A., & Tellegen, A. (1998). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54(6),* 1063-1070.

Westen, D., & Arkowitz-Westen, L. (1998). Limitations of Axis II in diagnosing personality pathology in clinical practice. *American Journal of Psychiatry, 155,* 1767–1771.

Westen, D., Bradley, R., & Hilsenroth, M. (2009).Unpublished manuscript, Emory University.

Westen, D., Heim, A.K., Morrison, K., Patterson, M., & Campbell, L. (2002). Simplifying diagnosis using a prototype-matching approach: Implications for the next edition of the DSM. In L.E. Beutler & M.L. Malik (Eds.), Rethinking The DSM: A Psychological Perspective (pp. 221–250). Washington, DC: American Psychological Association.

Westen, D., & Weinberger, J. (2004). When clinical description becomes statistical prediction. *American Psychologist, 59(7),* 595-613.

Westen, D., & Shedler, J. (1999). Revising and assessing Axis II, Part II: Toward an empirically based and clinically useful classification of personality disorders. *American Journal of Psychiatry, 156 (2)*, 273-285.

Westen, D., Shedler, J., & Bradley, R. (2006). A prototype approach to personality diagnosis. *American Journal of Psychiatry, 163*, 846-856.

Whisman, M.A., & Pinto, A. (1997). Hopelessness depression in depressed inpatient adolescents. *Cognitive Therapy and Research, 21,* 345-358.Widiger, T.A. (1982). Prototypic typology and borderline diagnoses. *Clinical Psychology Review, 2,* 115-135.

Widiger, T.A. (1983). Utilities and Fixed Diagnostic Rules: Comments on Finn. *Journal of Abnormal Psychology, 92(4),* 495-498.

Widiger, T.A. (1993). The DSM-III-R categorical personality disorders: A critique and an aleternative. *Psychological Inquiry, 4,* 75-90.

Widiger, T.A. (2000). Personality disorders in the 21st century. *Journal of Personality Disorders, 14,* 3-16.

Widiger, T.A., & Clark, L.A. (2000). Toward *DSM-V* and the classification of psychopathology. *Psychological Bulletin, 126,* 946-963.

Widiger, T.A., & Samuel, D.B. (2005). Diagnostic Categories or Dimensions? A Question for the *Diagnostic and Statistical Manual of Mental Disorders – Fifth Edition. Journal of Abnormal Psychology, 114(4),* 494-504.

Widiger, T.A., & Sanderson, C.J. (1995). Towards a dimensional model of personality disorders. In W.J. Livesley (Ed.), The DSM-IV personality disorders (pp. 433-458). New York: Guilford Press.

Widiger, T., & Simonsen, E. (2005). The American Psychiatric Association's research agenda for the DSM-V. *Journal of Personality Disorders, 19,* 103-109.

Table 1. *Predicted Associations between Diagnostic prototypes and Relevant Criterion Measures.*

	Major Depression Prototype	Dysthymia Prototype	Bipolar Prototype	Generalized Anxiety Prototype	Panic Prototype	PTSD Prototype	Anorexia Nervosa Prototype	Bulimia Nervosa Prototype
PAI Depression	++	+	+			+	+	+
PAI Depression-Cognitive subscale	++	+						
PAI Depression-Affective subscale	++	+						
PAI Depression-Physiological subscale	++	+						
PAI Suicidal Ideation	+	+						
Global Assessment of Functioning (GAF) scale score	-	-	--			-	-	
PAI Mania			++					
PANAS-Positive Affect	-	-		-	-	-	-	-
PANAS-Negative Affect	++	++		+	+	+	+	+
PAI Anxiety	+	+		++	++	++	+	+
PAI Anxiety-Cognitive subscale	+	+		++	+	+		
PAI Anxiety-Physiological subscale	+	+		+	++	++		
PAI Anxiety-Affective subscale	+	+		++	++	++		
PAI Anxiety Related Disorders	+	+		++	++	++	+	+
Anxiety Sensitivity Index				+	++			
PAI Traumatic Stress subscale						++		
EDDS-Symptom count							+	+
PAI Aggression								
PAI Alcohol								
PAI Drug								
PAI Antisocial	-			-				

+ small to moderate positive correlations, ++ moderate to large positive correlations, - small to moderate negative correlations, -- moderate to large negative correlations, blank cells indicate no predicated association.

Table 2. *Patient and Clinician Demographics and Characteristics*

Patient Characteristics				
	Mean	**SD**	**Min**	**Max**
Age	37.9	12.3	18	60
Global Assessment of Functioning	62.8	10.8	28	90
	N	**%**		
Sex				
Female	50	59.5		
Male	34	40.5		
Race				
Caucasian	66	79.5		
African-American	6	7.2		
Hispanic	1	1.2		
Asian	4	4.8		
Other	6	7.2		
Socioeconomic Status				
Poor	8	9.6		
Working class	21	25.3		
Middle class	35	42.2		
Upper middle class	17	20.5		
Upper class	2	2.4		
Clinician Characteristics				
	N	**%**		
Sex				
Female	40	47.6		
Male	25	29.8		
Unreported	19	22.6		
Discipline				
Psychiatry	20	24.1		
Psychology	46	55.4		
Social Work	17	20.5		

Site				
Emory Outpatient Psychotherapy Training Program	20	23.8		
Emory Psychological Center	24	28.6		
Grady Hospital	2	2.4		
Emory Psychopharmacology Clinic	1	1.2		
Cambridge Health Alliance	33	39.3		
Emory Counseling Center	1	1.2		
Unreported	3	3.6		
Year in Training				
2nd yr psychiatry	1	1.2		
3rd yr psychiatry	9	10.7		
4th yr psychiatry	11	13.1		
3rd yr phd	10	11.9		
4th yr phd	8	9.5		
5th yr phd	5	6.0		
6th yr phd	1	1.2		
Post-doc	20	23.8		
1st year social work	12	14.3		
2nd year social work	1	1.2		
Unreported	6	7.1		

Table 3. *Correlations between Diagnostic Prototypes and Relevant Criterion Measures.*

	Major Depression Prototype	Dysthymia Prototype	Bipolar Prototype	Generalized Anxiety Prototype	Panic Prototype	PTSD Prototype	Anorexia Nervosa Prototype	Bulimia Nervosa Prototype
PAI Depression	0.47**	0.37**	-0.02	0.12	0.11	0.07	-0.01	0.20
PAI Depression-Cognitive subscale	0.42**	0.28**	-0.02	0.11	0.12	0.08	-0.04	0.16
PAI Depression-Affective subscale	0.48**	0.36**	0.00	0.16	0.05	0.06	-0.01	0.12
PAI Depression-Physiological subscale	0.35**	0.32**	-0.04	0.04	0.11	0.07	0.02	0.24*
PAI Suicidal Ideation	0.42**	0.39**	0.09	0.12	-0.04	-0.04	-0.04	0.08
Global Assessment of Functioning (GAF) scale score	-0.42**	-0.33**	-0.20	-0.15	-0.11	-0.22*	-0.19	-0.25*
PAI Mania	0.12	-0.16	-0.06	0.04	-0.01	-0.05	0.01	-0.12
PANAS-Positive Affect	-0.34**	-0.27*	0.00	-0.18	0.05	0.01	0.03	-0.04
PANAS-Negative Affect	0.33**	0.24*	0.04	0.26*	0.26*	0.11	0.03	0.19
PAI Anxiety	0.27*	0.21	0.14	0.33**	0.37**	0.17	0.07	0.16
PAI Anxiety-Cognitive subscale	0.27*	0.20	0.12	0.35**	0.34**	0.16	0.08	0.13
PAI Anxiety-Physiological subscale	0.23*	0.22*	0.10	0.21	0.32**	0.11	0.06	0.15
PAI Anxiety-Affective subscale	0.23*	0.16	0.16	0.32**	0.35**	0.20	0.04	0.16
PAI Anxiety Related Disorders	0.24*	0.10	0.03	0.25*	0.25*	0.21	0.05	0.16
Anxiety Sensitivity Index	0.22*	0.03	-0.18	0.15	0.23*	0.11	-0.08	0.04
PAI Traumatic Stress Subscale	0.17	0.12	0.06	0.19	0.15	0.30**	0.04	0.14
EDDS-Symptom count	0.20	0.14	-0.06	0.08	0.11	0.10	0.11	0.39**
Discriminant Validity								
PAI Aggression	0.15	0.03	0.23*	0.07	-0.04	0.03	-0.05	0.07
PAI Alcohol Problems	0.07	0.07	-0.05	0.14	-0.06	-0.09	0.03	0.02
PAI Drug Problems	0.04	-0.01	0.21	0.09	-0.03	0.01	0.02	0.01
PAI Antisocial	-0.07	-0.07	-0.07	-0.11	-0.11	-0.11	-0.11	-0.11

*$p<0.05$,** $p<0.01$

Table 4: *Correlations between Categorical Mood Diagnoses and Relevant Criterion Measures*

	Categorical MDD dx	**Categorical Dysthymia dx**	**Categorical Bipolar I dx**
PAI Depression	0.50**	0.18	0.06
PAI Depression-Cognitive subscale	0.46**	0.13	0.19
PAI Depression-Affective subscale	0.50**	0.19	0.07
PAI Depression-Physiological subscale	0.35**	0.19	-0.09
PAI Suicidal Ideation	0.32**	0.21	0.11
Global Assessment of Functioning (GAF) scale score	-0.40**	-0.14	-0.17
PAI Mania	-0.06	-0.13	0.09
PANAS-Positive Affect	-0.44**	-0.16	0.01
PANAS-Negative Affect	0.22	0.01	0.08
PAI Anxiety	0.16	-0.07	0.20
PAI Anxiety-Cognitive subscale	0.14	-0.05	0.15
PAI Anxiety-Physiological subscale	0.14	0.00	0.12
PAI Anxiety-Affective subscale	0.15	-0.14	0.27*
PAI Anxiety Related Disorders	0.32**	-0.08	0.24*
Anxiety Sensitivity Index	0.29**	-0.09	0.08
PAI Traumatic Stress subscale	0.21	-0.03	0.19
EDDS-Symptom count	0.22*	0.07	-0.02
PAI Aggression	0.11	-0.04	0.12
PAI Alcohol Problems	0.06	0.14	-0.03
PAI Drug Problems	-0.08	0.01	-0.03
PAI Antisocial	-0.03	-0.07	0.06

*$p<0.05$,** $p<0.01$

Table 5: *Fisher's z Comparisons of Categorical versus Prototype Associations with Relevant Criterion Measures.*

	MDD		DYSTHYMIA		BIPOLAR	
	Fisher's z	Two-tailed p	Fisher's z	Two-tailed p	Fisher's z	Two-tailed p
PAI Depression	-0.40	0.69	**1.97**	**<0.05**	-0.61	0.54
PAI Depression-Cognitive subscale	-0.51	0.61	1.52	0.13	-1.62	0.11
PAI Depression-Affective subscale	-0.26	0.79	1.76	0.08	-0.53	0.59
PAI Depression-Physiological subscale	0.00	1.00	1.34	0.18	0.38	0.70
PAI Suicidal Ideation	1.23	0.22	1.89	0.06	-0.15	0.88
Global Assessment of Functioning (GAF) scale score	-0.25	0.80	-1.95	0.05	-0.23	0.82
PANAS-Positive Affect	1.24	0.21	-1.11	0.27	-0.08	0.94
PANAS-Negative Affect	1.30	0.19	**2.30**	**0.02**	-0.31	0.76
PAI Anxiety	1.28	0.20	**2.80**	**0.01**	-0.47	0.64
PAI Anxiety-Cognitive subscale	1.51	0.13	**2.49**	**0.01**	-0.23	0.82
PAI Anxiety-Physiological subscale	1.04	0.30	**2.20**	**0.03**	-0.15	0.88
PAI Anxiety-Affective subscale	0.92	0.36	**2.99**	**0.00**	-0.87	0.39
PAI Anxiety Related Disorders	-0.95	0.34	1.78	0.08	-1.63	0.10

Appendix A: ***Hierarchical Linear Regressions***

Prediction of PAI-Depression	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Categorical diagnosis for Major Depressive Disorder				.50	.25	26.20	.00
Major Depressive Disorder Categorical Diagnosis	.50	5.12	.00				
Step 2: Categorical diagnosis for Major Depressive Disorder plus Major Depressive Disorder prototype				.54	.29	4.70	.02
Major Depressive Disorder Categorical Diagnosis	.31	2.34	.01				
Major Depressive Disorder Prototype	.29	2.17	.02				
Prediction of PAI-Depression	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Prototype diagnosis for Major Depressive Disorder				.50	.25	25.25	.00
Major Depressive Disorder Prototype Diagnosis	.50	5.03	.00				
Step 2: Prototype diagnosis for Major Depressive Disorder plus Categorical diagnosis for Major Depressive Disorder				.54	.29	5.45	.01
Major Depressive Disorder Prototype Diagnosis	.29	2.17	.02				
Major Depressive Disorder Categorical Diagnosis	.31	2.34	.01				

Prediction of PAI-Depression-Cognitive	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Categorical diagnosis for Major Depressive Disorder				.47	.22	22.44	.00
Major Depressive Disorder Categorical Diagnosis	.47	4.74	.00				
Step 2: Categorical diagnosis for Major Depressive Disorder plus Major Depressive Disorder prototype				.50	.25	2.56	.06
Major Depressive Disorder Categorical Diagnosis	.32	2.39	.01				
Major Depressive Disorder Prototype	.22	1.60	.06				
Prediction of PAI-Depression-Cognitive	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Prototype diagnosis for Major Depressive Disorder				.44	.19	18.59	.00
Major Depressive Disorder Prototype Diagnosis	.44	4.31	.00				
Step 2: Prototype diagnosis for Major Depressive Disorder plus Categorical diagnosis for Major Depressive Disorder				.49	.25	5.73	.01
Major Depressive Disorder Prototype Diagnosis	.22	1.60	.06				
Major Depressive Disorder Categorical Diagnosis	.32	2.39	.01				

Prediction of PAI-Depression-Affective	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Categorical diagnosis for Major Depressive Disorder				.50	.25	25.92	.00
Major Depressive Disorder Categorical Diagnosis	.50	5.09	.00				
Step 2: Categorical diagnosis for Major Depressive Disorder plus Major Depressive Disorder prototype				.55	.30	5.18	.02
Major Depressive Disorder Categorical Diagnosis	.30	2.25	.02				
Major Depressive Disorder Prototype	.30	2.28	.02				
Prediction of PAI-Depression-Affective	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Prototype diagnosis for Major Depressive Disorder				.50	.25	26.05	.00
Major Depressive Disorder Prototype Diagnosis	.50	5.10	.00				
Step 2: Prototype diagnosis for Major Depressive Disorder plus Categorical diagnosis for Major Depressive Disorder				.55	.30	5.07	.02
Major Depressive Disorder Prototype Diagnosis	.30	2.28	.02				
Major Depressive Disorder Categorical Diagnosis	.30	2.25	.02				

Prediction of PAI-Depression-Physiological	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Categorical diagnosis for Major Depressive Disorder				.36	.13	11.29	.00
Major Depressive Disorder Categorical Diagnosis	.36	3.36	.00				
Step 2: Categorical diagnosis for Major Depressive Disorder plus Major Depressive Disorder prototype				.40	.16	2.70	.06
Major Depressive Disorder Categorical Diagnosis	.19	1.35	.09				
Major Depressive Disorder Prototype	.24	1.64	.06				
Prediction of PAI-Depression-Physiological	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Prototype diagnosis for Major Depressive Disorder				.37	.14	12.27	.00
Major Depressive Disorder Prototype Diagnosis	.37	3.50	.00				
Step 2: Prototype diagnosis for Major Depressive Disorder plus Categorical diagnosis for Major Depressive Disorder				.40	.16	1.83	.09
Major Depressive Disorder Prototype Diagnosis	.24	1.64	.06				
Major Depressive Disorder Categorical Diagnosis	.19	1.35	.09				

Prediction of Global Assessment of Functioning (GAF) Scale score	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Categorical diagnosis for Major Depressive Disorder				.37	.14	12.78	.00
Major Depressive Disorder Categorical Diagnosis	-.38	-3.57	.00				
Step 2: Categorical diagnosis for Major Depressive Disorder plus Major Depressive Disorder prototype				.43	.19	3.91	.03
Major Depressive Disorder Categorical Diagnosis	-.18	-1.26	.11				
Major Depressive Disorder Prototype	-.29	-1.98	.03				
Prediction of Global Assessment of Functioning (GAF) Scale score	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Prototype diagnosis for Major Depressive Disorder				.41	.17	15.47	.00
Major Depressive Disorder Prototype Diagnosis	-.41	-3.93	.00				
Step 2: Prototype diagnosis for Major Depressive Disorder plus Categorical diagnosis for Major Depressive Disorder				.43	.19	1.59	.11
Major Depressive Disorder Prototype Diagnosis	-.29	-1.98	.03				
Major Depressive Disorder Categorical Diagnosis	-.18	-1.26	.11				

Prediction of Composite Adaptive Functioning-Clinician Report	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Categorical diagnosis for Major Depressive Disorder				.34	.11	9.89	.00
Major Depressive Disorder Categorical Diagnosis	-.34	-3.14	.00				
Step 2: Categorical diagnosis for Major Depressive Disorder plus Major Depressive Disorder prototype				.35	.12	1.00	.16
Major Depressive Disorder Categorical Diagnosis	-.24	-1.61	.06				
Major Depressive Disorder Prototype	-.15	-1.00	.16				
Prediction of Composite Adaptive Functioning-Clinician Report	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Prototype diagnosis for Major Depressive Disorder				.31	.09	8.14	.01
Major Depressive Disorder Prototype Diagnosis	-.04	-2.85	.01				
Step 2: Prototype diagnosis for Major Depressive Disorder plus Categorical diagnosis for Major Depressive Disorder				.35	.12	2.59	.06
Major Depressive Disorder Prototype Diagnosis	-.15	-1.00	.16				
Major Depressive Disorder Categorical Diagnosis	-.24	-1.61	.06				

Prediction of Composite Adaptive Functioning-Patient Report	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Categorical diagnosis for Major Depressive Disorder				.38	.15	13.36	.00
Major Depressive Disorder Categorical Diagnosis	-.38	-3.66	.00				
Step 2: Categorical diagnosis for Major Depressive Disorder plus Major Depressive Disorder prototype				.39	.15	.38	.27
Major Depressive Disorder Categorical Diagnosis	-.32	-2.23	.02				
Major Depressive Disorder Prototype	-.09	-.62	.27				
Prediction of Composite Adaptive Functioning-Patient Report	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Prototype diagnosis for Major Depressive Disorder				.31	.10	8.23	.00
Major Depressive Disorder Prototype Diagnosis	-.31	-2.87	.01				
Step 2: Prototype diagnosis for Major Depressive Disorder plus Categorical diagnosis for Major Depressive Disorder				.39	.15	4.98	.02
Major Depressive Disorder Prototype Diagnosis	-.09	-.62	.27				
Major Depressive Disorder Categorical Diagnosis	-.32	-2.23	.02				

Prediction of PAI-Depression	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Categorical diagnosis for Dysthymic Disorder				.20	.04	3.15	.04
Dysthymic Disorder Categorical Diagnosis	.20	1.78	.04				
Step 2: Categorical diagnosis for Dysthymic Disorder plus Dysthymic Disorder prototype				.36	.13	7.91	.01
Dysthymic Disorder Categorical Diagnosis	-.01	-.11	.46				
Dysthymic Disorder Prototype	.37	2.81	.01				
Prediction of PAI-Depression	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Prototype diagnosis for Dysthymic Disorder				.36	.13	11.49	.00
Dysthymic Disorder Prototype Diagnosis	.36	3.39	.00				
Step 2: Prototype diagnosis for Dysthymic plus Categorical diagnosis for Dysthymic Disorder				.36	.13	.01	.46
Dysthymic Disorder Prototype Diagnosis	.37	2.81	.01				
Dysthymic Disorder Categorical Diagnosis	-.01	-.11	.46				

Prediction of PAI-Depression-Cognitive	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Categorical diagnosis for Dysthymic Disorder				.16	.03	2.08	.08
Dysthymic Disorder Categorical Diagnosis	.16	1.44	.08				
Step 2: Categorical diagnosis for Dysthymic Disorder plus Dysthymic Disorder prototype				.28	.08	4.30	.02
Dysthymic Disorder Categorical Diagnosis	.00	.01	.50				
Dysthymic Disorder Prototype	.28	2.07	.02				
Prediction of PAI-Depression-Cognitive	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Prototype diagnosis for Dysthymic Disorder				.28	.08	6.56	.01
Dysthymic Disorder Prototype Diagnosis	.28	2.56	.01				
Step 2: Prototype diagnosis for Dysthymic plus Categorical diagnosis for Dysthymic Disorder				.28	.08	.00	.50
Dysthymic Disorder Prototype Diagnosis	.28	2.07	.02				
Dysthymic Disorder Categorical Diagnosis	.00	.01	.50				

Prediction of PAI-Depression-Affective	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Categorical diagnosis for Dysthymic Disorder				.22	.05	3.96	.03
Dysthymic Disorder Categorical Diagnosis	.22	1.99	.03				
Step 2: Categorical diagnosis for Dysthymic Disorder plus Dysthymic Disorder prototype				.34	.11	5.63	.01
Dysthymic Disorder Categorical Diagnosis	.04	.31	.38				
Dysthymic Disorder Prototype	.31	2.37	.01				
Prediction of PAI-Depression-Affective	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Prototype diagnosis for Dysthymic Disorder				.34	.11	9.85	.00
Dysthymic Disorder Prototype Diagnosis	.34	3.14	.00				
Step 2: Prototype diagnosis for Dysthymic plus Categorical diagnosis for Dysthymic Disorder				.34	.11	.10	.38
Dysthymic Disorder Prototype Diagnosis	.31	2.37	.01				
Dysthymic Disorder Categorical Diagnosis	.04	.31	.38				

Prediction of PAI-Depression-Physiological	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Categorical diagnosis for Dysthymic Disorder				.14	.02	1.50	.11
Dysthymic Disorder Categorical Diagnosis	.14	1.23	.11				
Step 2: Categorical diagnosis for Dysthymic Disorder plus Dysthymic Disorder prototype				.34	.11	8.18	.01
Dysthymic Disorder Categorical Diagnosis	-.08	-.60	.28				
Dysthymic Disorder Prototype	.38	2.86	.01				
Prediction of PAI-Depression-Physiological	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Prototype diagnosis for Dysthymic Disorder				.33	.11	9.54	.00
Dysthymic Disorder Prototype Diagnosis	.33	3.09	.00				
Step 2: Prototype diagnosis for Dysthymic plus Categorical diagnosis for Dysthymic Disorder				.34	.11	.36	.28
Dysthymic Disorder Prototype Diagnosis	.38	2.86	.00				
Dysthymic Disorder Categorical Diagnosis	-.08	-.60	.28				

Prediction of Global Assessment of Functioning (GAF) Scale score	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Categorical diagnosis for Dysthymic Disorder				.14	.02	1.53	.11
Dysthymic Disorder Categorical Diagnosis	-.14	-1.24	.11				
Step 2: Categorical diagnosis for Dysthymic Disorder plus Dysthymic Disorder prototype				.31	.09	6.02	.01
Dysthymic Disorder Categorical Diagnosis	.04	.32	.38				
Dysthymic Disorder Prototype	-.33	-2.45	.01				
Prediction of Global Assessment of Functioning (GAF) Scale score	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Prototype diagnosis for Dysthymic Disorder				.30	.09	7.64	.01
Dysthymic Disorder Prototype Diagnosis	-.30	-2.76	.01				
Step 2: Prototype diagnosis for Dysthymic plus Categorical diagnosis for Dysthymic Disorder				.31	.09	.10	.38
Dysthymic Disorder Prototype Diagnosis	-.33	-2.45	.01				
Dysthymic Disorder Categorical Diagnosis	.04	.32	.38				

Prediction of Composite Adaptive Functioning-Clinician Report	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Categorical diagnosis for Dysthymic Disorder				.17	.03	2.27	.07
Dysthymic Disorder Categorical Diagnosis	-.17	-1.51	.07				
Step 2: Categorical diagnosis for Dysthymic Disorder plus Dysthymic Disorder prototype				.19	.04	.50	.24
Dysthymic Disorder Categorical Diagnosis	-.11	-.82	.21				
Dysthymic Disorder Prototype	-.10	-.71	.24				
Prediction of Composite Adaptive Functioning-Clinician Report	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Prototype diagnosis for Dysthymic Disorder				.16	.03	2.10	.08
Dysthymic Disorder Prototype Diagnosis	-.16	-1.45	.08				
Step 2: Prototype diagnosis for Dysthymic plus Categorical diagnosis for Dysthymic Disorder				.19	.04	.67	.21
Dysthymic Disorder Prototype Diagnosis	-.10	-.71	.24				
Dysthymic Disorder Categorical Diagnosis	-.11	-.82	.21				

Prediction of Composite Adaptive Functioning-Patient Report	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Categorical diagnosis for Dysthymic Disorder				.07	.01	.40	.27
Dysthymic Disorder Categorical Diagnosis	-.07	-.64	.27				
Step 2: Categorical diagnosis for Dysthymic Disorder plus Dysthymic Disorder prototype				.28	.08	5.88	.01
Dysthymic Disorder Categorical Diagnosis	.12	.86	.20				
Dysthymic Disorder Prototype	-.33	-2.43	.01				
Prediction of Composite Adaptive Functioning-Patient Report	*Stand.β*	*t*	*p*	*R*	R^2	*F* change	*p* change
Step 1: Prototype diagnosis for Dysthymic Disorder				.26	.07	5.59	.01
Dysthymic Disorder Prototype Diagnosis	-.26	-2.36	.01				
Step 2: Prototype diagnosis for Dysthymic plus Categorical diagnosis for Dysthymic Disorder				.28	.08	.74	.20
Dysthymic Disorder Prototype Diagnosis	-.33	-2.43	.01				
Dysthymic Disorder Categorical Diagnosis	.12	.86	.20				

Appendix B: *Prototype Matching Approach to Diagnosis*

Background We are testing an approach to clinical diagnosis aimed at to making the diagnostic manual more manageable and clinically useful. It resembles the familiar DSM-IV system but differs in some important ways.

First, the diagnostic categories are rated on a continuum, not just present/absent. Thus, you can diagnose patients who do not quite meet criteria for current disorders by indicating a moderate match between the patient and the prototype. For example, if a patient is clearly depressed but is not profoundly depressed enough to warrant a diagnosis of Major Depression, s/he would receive a moderate rating, indicating subclinical depression. This eliminates the need for many "not otherwise specified" (NOS) and subthreshold diagnoses.

Second, this system is designed to integrate diagnosis, case formulation, and treatment planning. If a patient has features of Major Depression, you code the severity of depressive thoughts and feelings, vegetative signs, and melancholic symptoms, as well as duration and age of onset, not just the global diagnosis. Finally, and most importantly, this system is designed to be clinician-friendly. You do not count symptoms or criteria, which clinicians tend not to do in practice. Rather, you simply rate the overall similarity or resemblance between your patient and a prototype that describes the disorder in its "purest" form.

Instructions What follows are prototypes of several disorders (including one personality disorder not included in DSM-IV). Read the description of each prototype, and form an *overall impression* of the disorder. When you have a good sense of the syndrome being described, rate the extent to which your patient matches (resembles) the prototype. Do not count symptoms or worry about whether individual statements apply. Instead, just consider the *overall* similarity between your patient and the prototype.

For each disorder, rate the patient using the 5-point rating scale shown below, by circling the appropriate number. Note that a rating of 4 or higher means the patient *has* the disorder. A rating of 3 means the patient does not reach threshold for a diagnosis but has significant *features* of the disorder.

1 little or no match (description does not apply)	
2 some match (patient has ***some features*** of this disorder)	
3 moderate match (patient has ***significant features*** of this disorder)	**Features**
4 good match (patient ***has*** this disorder; diagnosis applies)	**Diagnosis**
5 very good match (patient ***exemplifies*** this disorder; prototypical case)	

Dysthymic Disorder

Patients who match this prototype are characterized by chronically depressed mood over many years. The severity of their depression may fluctuate over time, but their depressed mood is enduring rather than episodic. They tend to have low self-esteem and to feel hopeless. They may have difficulty making decisions or concentrating at times. Patients who match this prototype may also have somatic symptoms of depression, such as low energy or fatigue, insomnia or hypersomnia, or poor appetite or overeating.

1 little or no match (description does not apply)	
2 some match (patient has ***some features*** of this disorder)	
3 moderate match (patient has ***significant features*** of this disorder)	**Features**
4 good match (patient ***has*** this disorder; diagnosis applies)	**Diagnosis**
5 very good match (patient ***exemplifies*** this disorder; prototypical case)	