Western Kentucky University
TopSCHOLAR®

8-1-2001

Literature Size Related to Diagnostic Inaccuracy of Personality Disorders

Vincent Intoccia
Western Kentucky University

Follow this and additional works at: http://digitalcommons.wku.edu/theses

 Part of the Psychology Commons

LITERATURE SIZE RELATED TO DIAGNOSTIC INACCURACY OF
PERSONALITY DISORDERS

A Thesis

Presented to

The Faculty of the Department of Psychology

Western Kentucky University

Bowling Green, Kentucky

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

Vincent A. Intoccia IV

August 2001

LITERATURE SIZE RELATED TO DIAGNOSTIC INACCURACY OF

PERSONALITY DISORDERS

Recommended _May 25, 2001_
(Date)

C. Clinton Layne
Director of Thesis

Pope Tarrence

Elmer Gray   _6/01/01_
Dean, Graduate Studies and Research   Date

Table of Contents

List of Illustrations

LITERATURE SIZE RELATED TO DIAGNOSTIC INACCURACY OF

PERSONALTIY DISORDERS

Vincent A. Intoccia IV   May 2001   28 pages

Directed by: C. Clinton Layne, John Bruni, and Jackie Pope

Department of Psychology   Western Kentucky University

Personality disorders offer clinicians a unique diagnostic challenge. The purpose of this study was to examine the relationship between literature size and diagnostic accuracy for personality disorders. The data used in this study were taken from the Blashfield & Intoccia (2000) and Blashfield & Herkov (1996) studies. The data were analyzed using a combination of correlations and single subject experimental designs. The results indicated that from 1980 to 1987 as literature size increased diagnostic accuracy increased across personality disorders. When examining literature growth three personality disorders (borderline, schizotypal, antisocial) appear to be carrying the literature growth for the group. These three personality disorders have a top five diagnostic accuracy rating. Possible explanations for this relationship as well as implications for future research are discussed.

Introduction

This study examines the relationship between literature size and diagnostic

accuracy for personality disorders. The personality disorders continue to represent a

specific diagnostic challenge for clinicians with everything from diagnostic model

debates to gender stereotypes seemingly playing a role in this challenge. In 1980 the

personality disorders were placed on axis (II) apart from most other disorders in the

DSM-III (American Psychiatric Association, Diagnostic and Statistical Manual of

Mental Disorders, 1980). In addition, specific inclusion and exclusion criteria were

added to each personality disorder diagnosis (Mellsop et al., 1982) to increase diagnostic

accuracy. Mellsop et al. (1982) found that there was poor inter-rater agreement for the

diagnosis of specific personality disorders with the best agreement for antisocial

(kappa=.49) and the worst agreement for schizoid (kappa=.01). In spite of the revisions

in the DSM-III, personality disorder diagnoses were still unreliable (Mellsop et al.,

1982).

The Diagnostic and Statistical Manual of Mental Disorders (DSM) uses a

categorical model of psychiatric classification, and that model is also widely used in the

medical field. The focus of this research is not dependent upon which model is used.

The focus will be on the current model employed in the DSM. In 1987 the DSM-III-R

(American Psychiatric Association, 1987) was published and substantial revisions were

made particularly to the descriptors of personality disorders. These revisions of axis II

were extensive because personality disorder diagnoses in general are less reliable than

axis I diagnoses (i.e. schizophrenia, Alzheimer's) and the coverage was seen as too

narrow, resulting in large numbers of diagnoses with little meaning (i.e. mixed

personality, atypical personality) (Morey, 1988). In general, in order to increase the reliability of a categorical diagnosis the coverage of the category must be decreased. If, however, the coverage is too narrow the result is large numbers of individuals with no diagnosis. This is the challenge of categorical models of classification. With respect to the DSM-III-R, coverage was increased, but at the cost of increased overlap among the disorders (Morey, 1988).

Yet another challenge with the personality disorders seems to be gender bias. Some disorders appear to represent common stereotypes in American culture. In a study conducted by Mellsop et al., (1982), the female patients received 73% and 88% of the diagnoses of histrionic and dependent personality disorders while the male patients received 67% of all the diagnoses of antisocial personality disorder. To put a twist on gender bias Morey & Ochoa (1989) found that the gender of clinicians emerged as a factor in the diagnosis of borderline personality disorder. Specifically, they found that female clinicians over diagnosed borderline personality disorder while male clinicians under diagnosed this disorder. There is an abundance of research dealing with gender bias of psychiatric disorders (see Rosenkrantz & Vogel 1970, Warner 1978, Slavney 1990).

Two important research studies were conducted examining the diagnostic accuracy of personality disorder, Morey & Ochoa (1989) using the DSM-III and Blashfield & Herkov (1996) study using the DSM-III-R. Results of the Blashfield & Herkov (1996) study were used in the present study to compare accuracy of diagnosis with size of the literature. This article was selected due to the greater similarity between the current diagnostic criteria in the DSM-IV-TR and the diagnostic criteria examined in

the study, selected from the DSM-III-R. In the Blashfield & Herkov (1996) study

clinicians were asked to select a client that they had diagnosed with a personality

disorder. This case selection would represent the "clinical-based" diagnosis. The

clinicians were then given a randomized list of 166 symptoms of the 11 personality

disorders included in the DSM-III-R, and were asked to check all of the symptoms that

applied to their patient. This would represent the "criterion-based" diagnosis.

Clinicians could err in one of two ways: they could either over diagnose a disorder

(symptoms were not sufficiently present) or under diagnose a disorder (symptoms

sufficiently present but diagnosis not assigned). The results were quite striking. The

mean agreement between the clinical and criterion based diagnosis was poor (overall

Kappa=.40). The clinicians' diagnoses did not agree with the DSM-III-R diagnoses in

60% of the cases. This figure represents a significant problem. Either American

clinicians don't function well in the role of diagnosticians or they simply do not follow

the mandated psychiatric classification system put forth in the DSM-III-R (Blashfield &

Herkov, 1996). Results of the Blashfield &Herkov (1996) were consistent with the

Morey & Ochoa (1989) study.

There have been numerous reasons for the diagnostic perils of the personality

disorders compared to other psychiatric diagnoses, but perhaps the most important issue

related to this group of disorders has yet to be touched upon; that is, the literature size of

these disorders. Scientific literature has been found to have an exponential growth

function. Specifically the scientific literature doubles once every 15 to 20 years (Price,

1963). Additionally, different scientific sub-fields grow at different rates. Sub-fields

that are considered "hot" attract new researchers, represent the latest technology, and have abundant funding (Menard, 1971).

When examining the limited research dealing with the sociology of the personality disorders a wealth of variation is discovered. Blashfield & Intoccia (2000) found that literature on "Personality Disorders" as a whole doubles about once every 20-25 years, consistent with general medical literature. Five of the disorders (dependent, narcissistic, obsessive-compulsive, paranoid, and passive-aggressive) have extremely small literatures, ten or fewer articles per year. Six disorders (dependent, histrionic, obsessive-compulsive, paranoid, passive-aggressive, and schizoid) had either flat (no appreciable slope) or negative (downward slope) growth slopes. Borderline personality disorder is the only disorder that is growing in literature size, and while antisocial personality disorder has a large literature, its growth as been stagnant for the past three decades (Blashfield & Intoccia, 2000).

When diagnostic criteria (labels) are formed it is from a pool of available literature. If that pool of literature is inadequate the diagnostic labels generated from that pool will be inadequate. The meaning is not that these labels are invalid. It simply indicates that without literature growth and the exploration of new information these labels become outdated and ineffective. Consider the following example. For many years general paresis was diagnosed using symptomatic indicators and this diagnostic system was accurate. However, due to continuing research on this disorder a pathogenic indicator the syphilis spirocii virus was discovered. Because of that discovery general paresis is now diagnosed earlier and more accurately.

The purpose of this study is to examine the diagnostic impact of literature size for each personality disorder to determine whether disorders with small or dying literatures are misdiagnosed to a greater extent than disorders with substantial or growing literatures. Specifically, the greater the literature base the greater diagnostic accuracy of personality disorder. The logic of this hypothesis is that growing literature or consistently high literature represents new empirical information and stimulates clinical interest in a personality disorder. With new empirical information comes the ability to refine the etiology and assessment of personality disorders.

Literature Review

There has been little research conducted on the sociology of science. In 1963 Derrick de Solla Price published a book entitled "Big Science Little Science." In this book Price asserts that scientific literatures from the 1600s to 1900s are doubling once every 15 years. He also found that scientific literature has an exponential growth function. To illustrate the magnitude of growth consider the following example: If in 1990 a university built a library that held a copy of every scientific publication from the beginning of time until 1990, this building would be a sizable one. If the university wanted to keep up with the growth rate it would have to build another building the same size to house the literature between 1990 and 2010. It would then have to build four buildings the same size to house the growth between 2010 and 2030. As illustrated, the growth of scientific literature is staggering.

Menard (1971) extended Price's work in a book entitled "Science Growth and Change." He began looking at the growth rate for different scientific subfields. He discovered different sub-fields grow at different rates. For example the literature on marine biology doubles about once every 5 years, whereas the literature on vertebrate paleontology doubles about once every 27 years. Menard (1971) then examined other aspects of different scientific sub-fields that might account for the differences in growth rates. It appears that certain aspects of a given sub-field greatly influence the growth of that field. For example, Menard (1971) stated that faster growing sub-fields have researchers whose average ages are younger, salaries are higher, and prestige is greater. Faster growing sub-fields represent the latest technology, have more available funding, and generate more interest in a given sub-field.

More recently Blashfield & Intoccia (2000) examined the sociology of personality disorders. The researchers examined the literature for each personality disorder from 1966 to 1995, by conducting a search using the Medline database indexing system. The initial finding was that as a whole the personality disorders literature appears to be growing at rate consistent with the general medical literature, doubling approximately once every 20-25 years. When examining the personality disorders at the individual level the results are quite different. Of the 11 personality disorders examined (taken from DSM-IV) three personality disorders appear to be carrying the growth for this group of disorders (borderline, schizotypal, and antisocial). Borderline personality disorder appears to be the only disorder generating substantial literature growth. Although, schizotypal personality is showing growth it has not reached over 100 articles per year at this time. In addition, antisocial personality disorder has a consistently high literature volume but has remained stagnant in growth (Blashfield & Intoccia, 2000).

Of the remaining disorders five appear to be "dead" meaning they have less than 10 articles a year published (avoidant, dependent, narcissistic, obsessive-compulsive, paranoid, and passive-aggressive) (Blashfield & Intoccia, 2000). There are seven disorders that are "flat" meaning they have no appreciable slope or are "dying" meaning downward growth slope. These disorders are dependent, avoidant, histrionic, obsessive-compulsive, paranoid, passive-aggressive, and schizoid. Another focus of this study was to examine the idea that literature growth was stimulated in 1980 when the personality disorders were placed on their own axis (II). Using the best fitting curve generated from a regression equation for the pre 1980 data, new points were generated post 1980. The conclusion was that the placement of the personality disorders on axis II did not

stimulate growth, evidenced by the post 1980 data falling beneath the regression line generated from the pre 1980 data (Blashfield & Intoccia, 2000).

In 1980 the publication of the DSM-III was the first attempt to provide specific and detailed information on personality disorder categories/criteria. With this new axis came new inclusion and exclusion criteria with the goal of increasing diagnostic accuracy (Mellsop et al., 1982). In a study conducted by Mellsop et al., (1982) three psychiatrists were instructed to diagnose the same twenty-five patients, and their blind ratings/diagnoses were then compared for level of agreement. The highest level of agreement was found for antisocial (k=.49), and the lowest for schizoid (k=.01). Mellsop et al. (1982) found that in spite of these inclusion and exclusion criteria, personality disorder diagnoses were still unreliable.

Morey (1988) examined the revisions of the DSM-III, know as the DSM-III-R. In the review the DSM-III-R underwent substantial revisions, with the most dramatic changes occurring on axis II. Axis II disorders were the focus of attention because they were found to be less reliable diagnostically than axis I disorders (Morey, 1988). Also, coverage for this group of disorders was considered too narrow, the result being a large number of patients with diagnoses that have little meaning (mixed or atypical personality). The DSM uses a categorical model of classification, so that reliability and coverage are inversely related. In order to increase diagnostic reliability for a specific disorder the coverage for that disorder would have to be decreased (Morey, 1988). As stated earlier, the axis II disorders in the DSM-III were lacking in both reliability and coverage. The DSM-III-R attempted to increase both the reliability and coverage of the personality disorders.

Morey & Ochoa's (1988) final data set consisted of 291 patients with a clinician

break down of 101 psychiatrist and 190 Ph.D. (clinical) psychologists. The clinicians

were randomly selected from the *Directory of Medical Specialist* and the *Directory of*

the American Psychological Association (Morey, 1988). The clinicians were instructed

to choose two patients currently meeting the criteria of any personality disorders. These

data were obtained by using the Vanderbilt Personality Disorders Diagnostic Survey.

This survey is a checklist of 166 criteria describing the 11 specific personality disorders.

These data were then calculated algorithmically using a computer program set to the

decision rules for the DSM-III and DSM-III-R (Morey, 1988). Morey (1988) found an

increase in coverage evidenced by the decrease in the diagnosis of mixed and atypical

personality from 29.2% to 22.3%. This outcome may have been accomplished by the

sharp rise in diagnostic overlap from 36.4% (meeting criteria for at least 2 PD) to 51.9%.

With regard to reliability comparisons the DSM-III and DSM-III-R diagnostic reliability

on a global level (mean level) were comparable. An increase in coverage was

accomplished, generally, without a decrease in diagnostic reliability. The cost of these

revisions was a dramatic increase in diagnostic overlap among the personality disorders

causing yet another challenge for clinicians.

Morey & Ochoa (1989) conducted an important study examining the diagnostic

accuracy of personality disorders. Implicit in their research is the idea that for the DSM-

III diagnostic strategy to be successful, clinicians must adhere to the specified diagnostic

criteria. The sample for the current study was taken from Morey (1988) described

above. Data were gathered from clinicians in a questionnaire format across the

following areas: demographic data (clients & therapists), clinical impression (multi-axial

diagnoses), and a randomized checklist of 166 symptoms of the 11 personality disorders in the DSM-III. The clinicians were asked to select one or more of their patients with a personality disorder diagnosis (multi-axial format) and describe their clinical impression of the patient. This impression became the "clinical diagnosis." The clinicians were then instructed to complete the checklist by checking all symptoms exhibited by the same client. The checklists were then analyzed algorithmically using a computer program guided by the decision rules specified in the DSM-III. This became the "criterion based diagnosis"(Morey & Ochoa, 1989).

Two types of diagnostic errors were analyzed, over-diagnosis and under-diagnosis. Overdiagnosis occurs when a diagnostic label is assigned and the sufficient criteria for that disorder are not met. Under-diagnosis occurs when the criteria for a specific diagnostic label are satisfied yet that diagnosis is not made. The following factors were examined as possible causes of bias in the present study: clinician experience (full-time years), clinician occupation (psychiatrist/psychologist), clinician gender, clinician orientation (psychodynamic/non-psychodynamic), clinical setting (outpatient/inpatient), patient age, patient gender, patient race (White/non-White), patient annual income, and patient education status (Morey & Ochoa, 1989). The following personality disorders were examined in detail due to their controversial categories: borderline, histrionic, antisocial, and dependent.

The results indicated that there was poor agreement between clinical and criterion based diagnosis with a mean agreement of k=.30 (Morey & Ochoa, 1989). The most salient predictor of both over-diagnosis and under-diagnosis of borderline personality was clinician experience. Specifically, inexperienced clinicians applied the label too

readily; conversely experienced clinicians did not assign the diagnosis when the client

did meet the DSM-III criteria (Morey & Ochoa, 1989). The second most salient

predictor of both over-diagnosis and under-diagnosis of borderline personality was

patient income, with wealthier clients being under-diagnosed and poorer clients being

over-diagnosed (Morey & Ochoa, 1989). The final significant demographic predictor

of both over-diagnosis and under-diagnosis of borderline personality was clinician

gender, with female clinicians tending to over-diagnose and male clinicians tending to

under-diagnose (Morey & Ochoa, 1989). Certain symptomatic predictors emerged as

possible causes of over-diagnosis and under-diagnosis of borderline personality. The

most salient symptomatic predictor of both over-diagnosis and under-diagnosis was

suicidal threats and gestures (Morey & Ochoa, 1989). The two most salient

symptomatic predictors of under-diagnosis of borderline personality were the absence of

suicidal threats and the presence of indifference to praise and criticism (Morey & Ochoa,

1989).

Morey & Ochoa (1989) found under-diagnosis of histrionic personality to be a

more common problem with this label. No demographic predictors were found to be

significant. The most powerful predictor of under-diagnosis of histrionic personality

was the presence of an identity disturbance and the absence of attention seeking (Morey

& Ochoa, 1989). When examining over-diagnosis of histrionic personality both

demographic and symptomatic predictors emerged as significant. The best symptomatic

predictors of over-diagnosis were "exaggerated expression of emotion" and "seductive

appearance or behavior" (Morey & Ochoa, 1989). The best demographic predictor of

over-diagnosis was patient age. The only significant clinician variable was gender, with male clinicians overly applying this label (Morey & Ochoa, 1989).

Antisocial personality was more reliably diagnosed than any other personality disorder, possibly due to its more objective criteria. The following four features appear to be the best predictors of over-diagnosis for this disorder: history of vandalism, exploitativeness, participation in illegal acts, and history of school expulsion (Morey & Ochoa, 1989). Due to the limited number of patients meeting the criteria for this disorder, a discriminate function model for under-diagnosis was not constructed (Morey & Ochoa, 1989).

The final disorder examined was dependent personality. When examining both over-diagnosis and under-diagnosis of dependent personality disorder, symptomatic features emerged as the best predictors. The best symptomatic predictors of over-diagnosis were: passively allowing others to assume responsibility, unjustified concern with fidelity of a spouse, lack of an identity disturbance, helplessness, and lack of success fantasies (Morey & Ochoa, 1989). One demographic predictor, clinician occupation, emerged significant, with psychologists tending to over-diagnose. The best symptomatic predictors of under-diagnosis were perfectionism, self-damaging impulsivity, avoidance of accepting blame, and inappropriate anger (Morey & Ochoa, 1989).

The results of this study indicate that clinicians do not closely follow the criteria set in the DSM-III when diagnosing personality disorders (Morey & Ochoa, 1989), as is evidenced by only a 30% mean agreement rate between clinical diagnoses and criterion diagnoses (Morey & Ochoa, 1989). Certain demographic features emerged as predictive

of over-diagnosis and under-diagnosis for specific personality disorders such as patient age, patient gender, patient income, clinician gender, clinician experience, and clinician occupation (psychiatrist/psychologist). In general it appears that symptomatic features were more predictive of under-diagnosis and over-diagnosis of personality disorders. Clinicians tend to place more weight on certain symptoms as indicative of specific disorders and tend not to utilize the additive feature model (decision rules) of the DSM-III (Morey & Ochoa, 1989).

In 1996 Blashfield & Herkov conducted a replication of the Morey & Ochoa (1989) study. Psychologists and psychiatrists were randomly selected using the yellow pages in the 50 most populous cities in the United States. Five hundred and twenty-two clinicians were selected. Approximately 62% completed all materials, representing 320 of their patients. Using a procedure similar to Morey & Ochoa (1989), clinicians were asked to select a patient with whom they had a minimum of 10 contact hours with and whom they believed had a personality disorder. The clinicians were then asked to state the patient's diagnoses. This clinical impression became known as the "clinical" diagnosis (clinicians were free to assign as many diagnoses they felt necessary and to specify primary or secondary status for a given diagnosis). Clinicians were then given a randomized list of the entire diagnostic criteria for all of the personality disorders listed in the DSM-III-R and asked to check all symptoms that applied to their client. The clinicians were not allowed to use the DSM-III-R to complete the checklist. The questionnaires were collected and the personality disorder criteria were then sorted into their appropriate diagnostic categories. The decision rules for the DSM-III-R were followed in order to obtain the "criterion-based" diagnosis. The agreement for the

clinical and DSM-III-R (criterion-based) diagnosis was obtained by using a kappa statistic.

There were two main differences in the current study and the Morey & Ochoa (1989) study. First, in the Morey & Ochoa study (1989) questions were asked about patient race, patient income, and clinician race; these questions were not included in the Blashfield & Herkov (1996) study. Another important difference was the basis of the criterion diagnosis. In the Morey & Ochoa (1989) study this diagnosis came from the DSM-III, while in the Blashfield & Herkov (1996) study the final version of the DSM-III-R was used. Blashfield & Herkov (1996) found support for the original Morey & Ochoa (1989) study with an overall mean agreement of k=.40, only k=.10 higher than the mean agreement for the Morey & Ochoa (1989) study. The highest level of agreement was k=.56 (borderline) and the lowest k=.28 (schizoid).

Three diagnostic criteria found to be predictive of over-diagnosis of borderline personality disorder were replicated (Blashfield & Herkov, 1996). They are as follows: recurrent suicidal gestures, affective instability, and self-damaging impulsivity. In reference to under-diagnosis of borderline personality disorder two criteria were replicated, absence of recurrent suicidal gestures and over concern with physical appearance (Blashfield & Herkov, 1996). For histrionic personality disorder two variables were replicated as significant predictors of over-diagnosis. These criteria were exaggerated expression of emotion and clinician gender, with male clinicians over-diagnosing (Blashfield & Herkov, 1996). Neither of the two variables in the Morey & Ochoa (1989) study that predicted under diagnosis of histrionic personality disorder were replicated in the current study (Blashfield & Herkov, 1996).

In regard to dependent personality disorder only one variable predictive of over-diagnosis in the Morey & Ochoa (1989) study was replicated: "Allows others to make most of his or her important decisions" (Blashfield & Herkov, 1996). Morey and Ochoa (1989) found five variables to be significant predictors of under-diagnosis of dependent personality disorder. These variables were not replicated nor were any new variables found to be predictive of under-diagnosis of dependent personality disorder at a significant level (Blashfield & Herkov, 1996). Roughly half of the significant predictors for over-diagnosis in the Morey & Ochoa (1989) study were replicated. Interestingly, practically none of the significant predictors of under-diagnosis in the Morey & Ochoa (1989) study were replicated. This result may be due to different statistical problems or the high predictor salience needed to cause under-diagnosis to occur (Blashfield & Herkov, 1996). The results of this replication found that in spite of numerous revisions to the DSM-III, the DSM-III-R did not increase diagnostic reliability.

There have been numerous factors identified that affect the diagnostic accuracy of personality disorders. The current study will examine another possible factor of diagnostic accuracy for the personality disorders. This factor is literature size for each disorder. This study will examine the relationship of empirical information to applied diagnostic labels. Specifically, personality disorders with large or growing literature bases will be diagnosed more accurately. The more empirical information that is available on a given disorder the more accurately that disorder can be diagnosed. This accuracy is most important because psychiatric diagnoses can have a profound impact on a client's personal and professional life. Clinicians must never forget without the "science" of psychology its professional application will fail, and it will fail clients.

There is no way to diagnose a psychiatric disorder reliably if only inadequate or outdated information is available on that specific disorder. With new empirical information comes the ability to refine the etiology and assessment of personality disorders.

Method

Data

The data used in this study were taken from the Blashfield & Intoccia (2000) and Blashfield & Herkov (1996) studies.

Subjects

In the Blashfield & Herkov (1996) study subjects were randomly selected using the yellow pages in the 50 most populous cities in the United States. Of the subjects selected there were 236 psychologists, 54 psychiatrists, and 25 other (e.g., MA). A total of five hundred twenty-two clinicians were selected. Approximately 62% completed all materials, representing diagnostic information on 320 of their patients.

Procedure

In the Blashfield & Herkov (1996) study clinicians were asked to select a patient they had a minimum of 10 contact hours with and whom they felt had a personality disorder. The clinicians were then asked to state the patient's diagnoses. This clinical impression became known as the "clinical" diagnosis (clinicians were free to assign as many diagnoses they felt necessary and to specify primary or secondary status for a given diagnosis). Clinicians were then given a randomized list of the entire diagnostic criteria for all of the personality disorders listed in the DSM-III-R and asked to check all symptoms that applied to their client. The clinicians were not allowed to use the DSM-III-R to complete the checklist. The questionnaires were collected and the personality disorder criteria were then sorted into their appropriate diagnostic categories. The decision rules for the DSM-III-R were followed in order to obtain the "criterion-based"

diagnosis. The agreement for the clinical and DSM-III-R (criterion-based) diagnosis was obtained by using a kappa statistic (see Table 1).

In the Blashfield & Intoccia (2000) study a comprehensive literature review from 1966 to 1995 was conducted using the Medline database. Annual data were gathered regarding the number of journal articles published per personality disorder, the total number of articles in Medline, and the total number of articles under the title "personality disorders." The data gathered under the title of "personality disorders" was actually the sum for each individual personality disorder. More specifically, because the search was done under the title "personality disorders" many unrelated articles were accessed. These data were first generated using the 1996, 1997, SilverPlatter International NV, WebSpirs version 3.1. They were later re-examined under the 1998 Gateway Ovid technologies. No significant variations in the number of articles per year were found as a function of the platforms used to access MEDLINE.

In the current study SPSS version 10 was used to combine and analyze the data. A combination of correlations and graphic representations were conducted. The literature base was split into two time periods 1966 to 1979 and 1980 to 1987. The data were split at 1980 because of the publication of the DSM-III in which the personality disorders were placed on their own axis (II). Correlations were generated across all years for all personality disorders. To carefully review at disorders in a more detailed way, data were analyzed as an N=1 study. This analysis included the mean and standard deviation for each personality disorder pre and post 1980.

Table 1

Diagnostic Accuracy Ratings for each Personality Disorder

	Clinical Diagnosis	Criterion Diagnosis	Current Kappa	Morey Kappa
	n =	n =		
Paranoid	41	60	.31	.31
Schizoid	22	57	.28	.11
Schizotypal	11	20	.56	.12
Antisocial	24	20	.46	.53
Borderline	102	85	.63	.59
Histrionic	39	65	.30	.27
Narcissistic	64	69	.54	.31
Avoidant	35	47	.30	.23
Dependent	61	57	.29	.23
Obsessive-compulsive	60	38	.38	.34
Passive-aggressive	51	66	.31	.28
Mean kappa value			.40	.30

Note. Adapted from "Investigation clinician adherence to

diagnosis by criteria: A replication of Morey and Ochoa (1989),"

by R.K. Blashfield, & J. Herkov, 1996, Journal of Personality

Disorders, 10(3), p. 223.

Results

The data on literature size were split into two time periods 1966-1979 and 1980-1987. A mean was then computed for each period. The mean for the literature size across disorders between 1966-1979 was 18 articles per year ($\underline{M}$1966 to1979= 18), and the mean for the literature size across disorders between 1980-1987 was 29 articles per year ($\underline{M}$1980 to 1987= 29). When examining this split the following personality disorders show literature growth: borderline, narcissistic, schizotypal, and paranoid. Although antisocial personality disorder does not appear to be growing there is a consistent and substantially high volume of literature being generated on this personality disorder.

In order to examine the relationship between literature size and diagnostic accuracy across disorders a Pearson's Correlation was used. The relationship between literature size and diagnostic accuracy for period 1 (1966 to 1979) was not significant p=.80 . For period 2 (1980 to 1987) a significant correlation was found (p=.03) with a Pearson's correlation of $\underline{r}$ =.672. Evidenced by these results for the years 1980 to 1987, as literature size increases diagnostic accuracy increases across personality disorders.

Perhaps the most effective way to analyze the relationship between literature size and diagnostic accuracy on an individual basis is to use a logic driven approach and examine the personality disorders with the top five diagnostic accuracy ratings. The most effective way to conduct this analysis is by using an N=1 study which treats each personality disorder as its own experiment (see Table 2).

Borderline personality disorder has the highest diagnostic accuracy rating (k= .63) and shows dramatic growth with a $\underline{M}$1966 to 1979 = .1429 and a $\underline{M}$1980 to1987 =

101.3. Schizotypal personality disorder has the second highest level of diagnostic agreement (k=.56) and shows strong literature growth with a M1966 to 1979= 13.7 and M1980 to 1987= 38.7. Narcissistic personality disorder has the third highest level of diagnostic agreement (k= .54) and shows a four fold increase in literature with a M1966 to 1979 =1.92 and a M1980 to 1987 = 7.75. The disorder with the fourth highest level of diagnostic accuracy (k= .46) is antisocial personality disorder. Although the literature base for antisocial personality disorder is not growing, it is marked by a consistently large number of articles per year (M1966 to 1979 =111.7 and M1980 to 1987 =96.3). Compulsive personality disorder has the fifth highest level of diagnostic accuracy (k=.38) but shows a decrease in literature size (M1966 to 1979 = 9 and M1980 to 1987 = 7.7).

Table 2

N=1 Single Subject Design Data Split Pre and Post 1980

Personality Disorders	Split	N # of years	Mean # of Articles	Std. Deviation	Diagnostic Accuracy
Antisocial	**.00**	**14**	**111.714**	**12.461**	**.46**
	1.00	**8**	**96.375**	**14.151**	
Schizotypal	**.00**	**14**	**13.785**	**15.166**	**.56**
	1.00	**8**	**38.750**	**14.469**	
Borderline	**.00**	**14**	**.1429**	**.3631**	**.63**
	1.00	**8**	**101.375**	**31.919**	
Schizoid	.00	14	15.741	14.090	.28
	1.00	8	12.250	3.105	
Histrionic	.00	14	16.571	8.317	.30
	1.00	8	15.625	3.814	
Compulsive	**.00**	**14**	**9.071**	**11.118**	**.38**
	1.00	**8**	**7.750**	**3.058**	
Narcissistic	**.00**	**14**	**1.928**	**2.585**	**.54**
	1.00	**8**	**7.750**	**3.535**	
Dependent	.00	14	2.714	4.322	.29
	1.00	8	4.500	3.338	
Passive-Aggressive	.00	14	2.714	3.646	.31
	1.00	8	1.750	1.669	
Paranoid	.00	14	3.000	3.637	.31
	1.00	8	6.250	3.494	

Note. The personality disorders in bold print are the disorders with the top five levels of diagnostic accuracy.

Discussion

This study was designed to examine another possible explanation for the diagnostic challenges of the personality disorders. Implicit in this research is the idea that growing literature bases allow for the refinement of diagnostic categories and more accurate diagnoses of personality disorders. Therefore, when examining the literature bases for the personality disorders one would expect to find that the greater the literature base, the greater the diagnostic accuracy for the personality disorders. The results of the current study support this hypothesis. Between the years 1966-1979 no significant relationship (r =.09, p=.8) existed between diagnostic accuracy and the size of the literature base. For the years 1980-1987 a significant correlation was found (r =.672, p=.03) between diagnostic accuracy and size of literature base. As literature size increases so does diagnostic accuracy across disorders from 1980-1987.

A single subject analysis was conducted in order to examine the relationship between diagnostic accuracy and literature growth for individual personality disorders. Borderline, schizotypal, and antisocial all have literature bases that are classified as "alive" or growing. These three disorders fit very neatly into the predictions of this research. They comprise the largest and healthiest literature bases and all have high diagnostic accuracy ratings. Narcissistic (k=.54) and compulsive (k=.38) personality disorders have the third and fifth highest diagnostic accuracy ratings but do not have substantial literature bases. Interestingly, narcissistic personality disorder is considered to have a "dead" literature base (less than ten articles per year) but shows dramatic growth with a four-fold increase in literature. Even though the literature base for narcissistic personality disorder is considered inadequate the undeniable growth of this

literature may help explain its high diagnostic accuracy. Compulsive personality disorder shows mild decline and is considered to have a "dead" literature base yet has the fifth highest diagnostic accuracy rating. One possible explanation for the high diagnostic accuracy of compulsive personality disorder is the number of behavioral indicators in its diagnostic criteria. For example, "preoccupied with rules, details, lists, order, organization, or schedules to the extent that the major point of the activity is lost; shows perfectionism that interferes with task completion; is excessively devoted to work and productivity to the exclusion of leisure activities and friendships."

The importance of this research is best seen when examining the possible circular pattern of research and clinical practice. Personality disorders that are stimulating clinical interest are likely stimulating research (Blashfield & Intoccia, 2000). With a growing and changing pool of literature more specific diagnostic features may be formed. For example the DSM-II (1968) defines obsessive compulsive personality as a behavior pattern characterized by excessive concern with conformity and adherence to standards of conscience. Individuals may be rigid, over-inhibited, over-conscientious, over-dutiful, and unable to relax. This description is a valid yet surprisingly muddy picture of how an individual with this personality disorder might appear. In the DSM-IV (1994) obsessive compulsive personality disorder is describe as a pervasive pattern of preoccupation with orderliness, perfectionism, and mental and interpersonal control, at the expense of flexibility, openness, and efficiency characterized by four of the following: "1) preoccupied with rules, lists, order, details, organization, or schedules to the extent that the major point of the activity is lost, 2) shows perfectionism that interferes with task completion,3) is excessively devoted to work and productivity to the

exclusion of leisure activities and friendships,4) is overconscientious, scrupulous, and inflexible about matters of morality, ethics, of values, 5) is unable to discard worn-out worthless objects even when they have no sentimental value, 6) is reluctant to delegate tasks or to work with others unless they submit to exactly his or her way of doing things, 7) adopts a miserly spending style toward both self and others, 8) shows rigidity and stubbornness." The diagnostic features in the DSM-IV described above offer a much clearer picture of the same personality disorder. The general definition of obsessive compulsive personality in the DSM-IV is not only much more detailed, but the diagnostic features are mostly behavioral indicators, making it more likely to diagnose this personality disorder accurately. Specific diagnostic features lead to more accurate diagnostic categories. As a result the more precisely a personality disorder is defined diagnostically the more specific the research questions generated for that disorder. One could argue that borderline, schizotypal, and antisocial are all stimulating clinical interest. Therefore, substantial research is being generated on these disorders. As a result the diagnostic categories of these disorders appear to be more clearly defined, evidenced by their high diagnostic accuracy.

It can also be argued that this possible circular pattern likely emerged in 1980 with the publication of the DSM-III. In the DSM-I and II personality disorders were simply described in a prose fashion typically with one paragraph and lists of descriptors but with no categorical structure. With the publication of the DSM-III significant revisions occurred across this group of disorders. Each disorder was described as a diagnostic category (categorical model) in an additive feature model. Specific information on particular personality features was given, and if a specified number of

features exist the diagnosis was warranted. Blashfield & Intoccia (2000) found that placing the personality disorders on axis II in 1980 did not stimulate literature growth for this group of disorders. This finding provided researchers with much more specific descriptions of each of the personality disorders. With these new more specific descriptions researchers were likely to generate more specific experimental questions, thereby offering one explanation for the significant correlation between diagnostic accuracy and literature size for the years 1980-1987.

When examining this study for limitations two main factors appear. First, the diagnostic accuracy ratings were taken from the DSM-III-R. It would be more beneficial if diagnostic accuracy ratings were obtained from the most recent diagnostic manual the DSM-IV-TR. Second, because the DSM-III-R was used the literature base was cut off the year it was published. The most striking results came from the second time period (1980-1987). By extending the years in the second time period one would expect to find greater support for the relationship between literature size and diagnostic accuracy. In order to provide the most up-to-date diagnostic accuracy ratings future research should start with a replication of the Morey & Ochoa (1989) study using the DSM-IV-TR (American Psychiatric Association, 2000). Using the DSM-IV-TR and expanding the literature base would provide the most up-to-date information on the relationship between diagnostic accuracy and literature size for the personality disorders.

As stated earlier in this study numerous factors have been found to affect the diagnosis of the personality disorders, including clinician gender, patient gender, patient income, patient education, and clinician occupation. The purpose of this study was to provide another possible factor for the diagnostic challenges of the personality disorders.

That factor is the literature base for the personality disorders. Empirical literature is perhaps the most important factor when examining this set of disorders. When personality disorders with inadequate empirical information are defined for diagnostic categories those categories are likely inadequate. With future research on the diagnostic factors affecting the personality disorders as well as the disorders themselves clinicians may more accurately diagnose and treat this perplexing set of disorders.

References

American Psychiatric Association. (1968). Diagnostic and Statistical Manual of Mental Disorders (2nd ed.). Washington D.C.: Author

American Psychiatric Association. (1980). Diagnostic and Statistical Manual of Mental Disorders (3rd ed.). Washington D.C.: Author.

American Psychiatric Association. (1987). Diagnostic and Statistical Manual of Mental Disorders (3rd ed. revised). Washington D.C.: Author.

American Psychiatric Association. (2000). Diagnostic and Statistical Manual of Mental Disorders (4th ed text revision.). Washington D.C.: Author.

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

Blashfield R.K., & Intoccia V (2000). Personality disorders growth and change. *American Journal of Psychiatry*, 157: 472-473.

Mellsop, C., Varghese, F., Joshua, S., & Hicks, A. (1982). The reliability of Axis II of DSM-III. *American Journal of Psychiatry*, 139, 1360-1361.

Menard H.W. (1971). Science: Growth and Change. Cambridge, MA: Harvard University Press.

Morey, L. C. (1988a). Personality disorders in DSM-III and DSM-III-R: Convergence, coverage, and internal consistency. *American Journal of Psychiatry*, 145(5), 573-577.

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

Price D.S. (1963). Big Science , Little Science. New Haven, Conn: Yale University Press.

Rosenkraz, P.S., & Vogel, S.R. (1970). Sex-Role Stereotypes and clinical judgments of mental health. *Journal of Consulting and Clinical Psychology*, 34(1), 1-7.

Slavney, P. R., (1990). Perspectives on "hysteria." Baltimore: Johns Hopkins University Press.

Warner, R., (1978). The diagnoses of antisocial and hysterical personality disorders: an example of sex bias. *Journal of Nervous Mental Disorder*, 166: 839-845.