

8-1-2008

Review of diagnostic methods in the most cited articles for anorexia nervosa, bulimia nervosa, and binge-eating disorder

Rebecca Chapman Roberts

Follow this and additional works at: <https://scholarsjunction.msstate.edu/td>

Recommended Citation

Roberts, Rebecca Chapman, "Review of diagnostic methods in the most cited articles for anorexia nervosa, bulimia nervosa, and binge-eating disorder" (2008). *Theses and Dissertations*. 3815.
<https://scholarsjunction.msstate.edu/td/3815>

This Graduate Thesis - Open Access is brought to you for free and open access by the Theses and Dissertations at Scholars Junction. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholars Junction. For more information, please contact scholcomm@msstate.libanswers.com.

REVIEW OF DIAGNOSTIC METHODS IN THE MOST CITED ARTICLES
FOR ANOREXIA NERVOSA, BULIMIA NERVOSA, AND
BINGE-EATING DISORDER

By

Rebecca Chapman Roberts

A Thesis
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in Psychology
in the Department of Psychology

Mississippi State, Mississippi

August 2008

REVIEW OF DIAGNOSTIC METHODS IN THE MOST CITED ARTICLES
FOR ANOREXIA NERVOSA, BULIMIA NERVOSA, AND
BINGE-EATING DISORDER

By

Rebecca Chapman Roberts

Approved:

Kevin J. Armstrong
Graduate Coordinator of the
Department of Psychology
(Director of Thesis)

Stephen B. Klein
Department Head of the
Department of Psychology

Kristine M. Jacquin
Associate Professor of Psychology
(Committee Member)

Deborah K. Eakin
Assistant Professor of Psychology
(Committee Member)

Gary L. Myers
Dean of the
College of Arts and Sciences

Name: Rebecca Chapman Roberts

Date of Degree: August 9, 2008

Institution: Mississippi State University

Major Field: Psychology (Clinical)

Major Professor: Dr. Kevin J. Armstrong

Title of Study: REVIEW OF DIAGNOSTIC METHODS IN THE MOST CITED
ARTICLES FOR ANOREXIA NERVOSA, BULIMIA NERVOSA,
AND BINGE-EATING DISORDER

Pages in Study: 111

Candidate for Degree of Master of Science

Diagnostic practices utilized in studies of participants with Anorexia Nervosa, Bulimia Nervosa, and Binge-Eating Disorder were investigated. A reliable coding system was used by two reviewers to analyze the diagnostic documentation practices in articles from the top-thirty most cited articles for each diagnostic category. Interrater agreements were all above .95. Results showed that many important diagnostic practices and criteria are either not being employed or not being documented. Uniform reporting procedures are necessary to help readers know how each article's sampling procedure and subject pool differs from other samples used in the literature. Researchers reported the most details about sample characteristics with the recently proposed Binge-Eating Disorder category. Discussion focuses on identifying the specific diagnostic and sampling procedures deserving better documentation in the eating disorder literatures.

DEDICATION

I would like to dedicate this paper to my parents, Paul and Nancy Roberts, without whose support, financially and emotionally, my pursuit of a Master's degree would not have been possible.

ACKNOWLEDGEMENTS

I would like to express my sincerest appreciation to the following people. First and foremost, I would like to thank Dr. Kevin Armstrong, my committee chair, for the opportunities he has provided me during my master's program, the faith he has had in me, and the time and effort he has given to me while completing my thesis. Second, I would like to thank Dr. Kristine Jacquin and Dr. Deborah Eakin, my committee members, for their support and guidance.

TABLE OF CONTENTS

DEDICATION	ii
ACKNOWLEDGEMENTS	iii
LIST OF TABLES	vii
CHAPTER	
I. INTRODUCTION	1
Diagnostic Controversies and Inconsistencies	4
Sample Characteristics	4
Diagnostic Methods	8
Self-report questionnaires	8
Food diaries	8
Semi-structured and structured interviews	9
Definition of Diagnostic Criteria	10
Body weight	10
Binge	11
Compensatory behavior	12
Adherence to Criterion Thresholds	13
Amenorrhea	13
Body weight	14
Binge frequency	14
Compensatory behavior frequency	16
Purpose of Study and Expected Findings	16
II. METHODS	19
Materials	19
Procedure	20

III. RESULTS	22
Overview of Analyses.....	22
Types of Studies Review	24
Study Type	24
Investigatory Department.....	25
Journal.....	25
Year.....	27
Times Cited.....	28
Demographics of Eating Disorder Samples.....	28
Age.....	28
Gender.....	30
Race.....	32
Socio-Economic Status	35
Referral Source	35
Subtypes.....	36
Anorexia Nervosa subtypes	37
Bulimia Nervosa subtypes	37
Eating Disorder Duration.....	37
History of Other Eating Disorder.....	38
Dual Diagnoses	39
Exclusionary Criteria	40
Diagnostic Methods Used.....	41
Diagnostic Methods Reported.....	41
Diagnoses Confirmed by Researchers	41
Self-Report Questionnaires.....	42
Food Diaries.....	43
Interviews.....	43
Structured interviews	43
Semi-structured interviews	43
Definition of Diagnostic Criteria	45
Definition of Body Weight	45
Definition of Binge	45
Definition of Compensatory Behavior.....	48
Adherence to DSM Criteria	50
Adherence to Amenorrhea Criterion.....	50
Adherence to Body Weight Criterion	51
Adherence to Binge Frequency Criterion	51
Adherence to Compensatory Behavior Frequency Criterion.....	53

IV. DISCUSSION.....	55
Types of Studies Reviewed.....	55
Demographics Reported Across Eating Disorder Samples	56
Establishment of Eating Disorder Diagnosis.	60
Definition of Diagnostic Criteria	62
Criterion Adherence.....	64
Summary of AN, BN, and BED Diagnostic Procedures	66
Limitations of the Current Research.....	68
Recommendations for Documentation of Diagnostic Practices	70
REFERENCES	72
APPENDIX	
A. ANOREXIA NERVOSA CODING SHEET.....	81
B. BULIMIA NERVOSA CODING SHEET.....	92
C. BINGE-EATING DISORDER CODING SHEET	103
D. IBR WAIVER – NONHUMAN SUBJECTS STUDY.....	111

LIST OF TABLES

1	Types of Studies Reviewed Overall.....	24
2	Investigatory Department Responsible	25
3	Journal Sources of Top Cited Articles	26
4	Publication Years of Top Cited Articles.....	27
5	Number of Citations for Top Cited Articles	28
6	Reporting of Age Demographics	29
7	Age Demographics Reported.....	30
8	Gender Demographics When Gender was Reported	32
9	Race Demographics Reported.....	34
10	Referral Populations Reported.....	36
11	Treatment of History of Other Eating Disorder	39
12	Dual Diagnoses Allowed in Eating Disorder Populations.....	40
13	Exclusionary Diagnoses Mentioned	41
14	Diagnosis Confirmed By Researchers	42
15	Interview Used.....	44
16	Reporting Binge Definition.....	46
17	Binge Definitions Reported	47

18	Reporting Diagnostic Definition of Compensatory Behavior	48
19	Compensatory Behavior Definitions Reported.....	50
20	Reporting Binge Frequency Requirements.....	52
21	Reporting Compensatory Behavior Frequency Requirements	53
22	Compensatory Behavior Frequency Cut-Offs Mentioned	54

CHAPTER I

INTRODUCTION

The diagnostic criteria for eating disorders have received considerable attention over the past two decades (Williamson, Zucker, Martin, & Smeets, 2001). Since their introduction in the *Diagnostic and Statistical Manual* (APA, 1980), the growing eating disorder literature has been rife with conflicting findings and inconsistent results. With subsequent revisions of the *DSM*, the eating disorders Anorexia Nervosa (AN) and Bulimia Nervosa (BN) received considerable modification to the core diagnostic sets (Devlin, Goldfein, & Dobrow, 2003), however, improvements in the specifications of criteria sets have yet to translate into consistent findings within the literature (Herpertz-Dahlmann, Muller, Herpertz, & Heussen, 2001). Researchers continue to struggle to integrate the expanding literature, while significant developments are potentially delayed as a result (Hsu, 1980; Strober, Freeman, & Morrell, 1997).

Concern regarding how eating disorder samples are recruited, diagnosed and defined has been highlighted as a potential explanation for the incohesive AN and BN literature bases (Eddy, Keel, & Dorer, 2002; Godart, Flament, Perdereau, & Jeammet, 2002; Klump, Bulik, & Pollice, 2000; Spindler & Milos, 2007; Sullivan, Bulik, Fear, & Pickering, 1998). Many researchers disagree about the core symptoms, the definition of criteria, and the conceptualization of these eating disorders (Garfinkel, Kennedy, &

Kaplan, 1996; Spoor, Stice, Burton, & Bohon, 2007; Sullivan, Bulik, & Kendler, 1998; Swain, Shisslak, & Crago, 1991; Wilson & Eldredge, 1991). Therefore, diagnostic approaches may vary from one study to another and introduce heterogeneity within samples across the literature. Unnecessary heterogeneity is a significant problem for any literature base and often results from inconsistency in the utilization of standard diagnostic procedures (Kazdin, 1995). For example, research in the area of Attention Deficit-Hyperactivity Disorder is largely muddled by the inconsistent diagnostic procedures and criteria thresholds utilized by clinical researchers (Myers, 2002) and across research disciplines (Dawkins, 2004; Hartley, 2003). Unnecessary heterogeneity is mainly introduced into research in two ways: (1) when investigations are incompatible in their definition of key criteria, and (2) when diagnostic thresholds are inconsistently maintained. Inconsistencies in these areas may result in divergent inclusionary and exclusionary decisions regarding research samples.

For the eating disorders Anorexia Nervosa (AN) and Bulimia Nervosa (BN), which have been criticized for being too diagnostically rigid and exclusive (the majority of eating disorder patients are relegated to the eating disorder not otherwise specified category based on inability to meet key criteria thresholds), maintaining a firm inclusionary boundary is necessary in evaluating the validity of such criticism. Researchers who waver in this regard may be including patient variables that heterogenize the research across samples, possibly resulting in divergent outcome data. However, in consideration of the evolving AN and BN literature, clear documentation of diagnostic procedures would allow for research to retain value over time. Failure to clearly detail sample in research may frustrate clinicians hoping to clarify and utilize

research contributions. Regarding this point, it is unclear as to the degree research use of diagnostic sampling procedures, diagnostic definitions, and diagnostic thresholds are inconsistent across the eating disorder literature for Anorexia and Bulimia Nervosa.

In the most recent revision of the *DSM* (APA, 2000), Binge-Eating Disorder (BED) emerged with provisional diagnostic status. Though not officially recognized as an independent diagnostic category, BED has received significant amounts of research attention. In the literature, BED is alternately conceptualized as a distinct disorder that differs qualitatively from other eating disturbances (Williamson, et al., 2002), as a variant of bulimia nervosa (Joiner, Vohs, & Heatheron, 2000), as a behavioral subtype of obesity (Devlin et al., 2003), or even as an associated feature of a separate primary disorder (Stunkard & Allison, 2003). However, the level of impairment associated with BED, in terms of accompanying distress (Devlin et al., 2003), body image dissatisfaction (Grilo, Masheb, & Wilson, 2001), and psychopathology (de Zwaan, Mitchell, & Seim, 1994; Masheb & Grilo, 2000; Wilfley, Schwartz, Spurrell, & Fairburn, 2000), have found to be comparable to other clinically significant eating disturbances. Distinctions have been demonstrated between BED and the non-binging obese (Guss, Kissileff, Devlin, Zimmerli, & Walsh, 2002), and bulimics (Striegel-Moore, Wilson, & Wilfley, 1998). The clinical presentation of uncompensated binge eating does exist, and provides significant impairment in functioning for those individuals who warrant the diagnosis. In light of the controversy surrounding conceptualization of BED, further systematic research is necessary to more clearly identify the significance of the BED phenomenon and to guide more effective treatment and prevention approaches (Wilfley et al., 2000). In particular,

further research will no doubt be needed regarding the implications of using inconsistent diagnostic approaches for implementing inclusionary/exclusionary criteria.

Diagnostic Controversies and Inconsistencies

To create a cohesive literature base, researchers must be consistent in diagnostic sampling procedures (Kazdin, 1995). Research can be consistent and descriptive regarding diagnostic procedures in three key ways, and by being so, researchers will greatly increase the comparability across the literature.

First, researchers investigating eating disorders minimally should facilitate comparability by documenting sampling procedures, including sample characteristics and diagnostic methods. Secondly, the definition of controversial criteria necessary for diagnosis should be provided because divergent definitions may introduce heterogeneity into research samples. Third, researchers ideally should supply valuable diagnostic threshold information regarding their diagnostic methodologies. The consistent use of key diagnostic criteria and accompanying thresholds allows for research to be adequately evaluated. Minus such information, clinicians and consumers are left to assumptions regarding the eating disorder profiles represented in the literature.

Sample Characteristics

Researchers may draw samples from previously diagnosed inpatient or outpatient populations or from the community. However, the majority of eating disordered individuals do not seek medical attention and may differ greatly from clinical populations typically represented in eating disorder research (Fairburn & Cooper, 1982; Fairburn, Welch, Doll, Davies, & O'Connor, 1997). For instance, depending on whether the eating

disordered participants are inpatient or outpatient, there can be considerable variation in the instance of comorbidity, and severity of eating disorder symptomology (Godart et al., 2002). This variability is expected in research, as different investigators conduct research with different goals in mind. However, different referral populations contain participants with starkly different presentations and eating disturbances, and therefore their inclusion in the data should be clearly documented (Welch & Fairburn, 1994).

Inclusionary and exclusionary decisions based on age, sex, race, socio-economic status, diagnostic subtypes, duration of the eating disorder, history of other eating disorder(s), and comorbid axis I or II disorders are useful for investigators in many cases. However, many of these variables correlate strongly with particular pathological profiles.

For example, sample characteristics of age and gender differ across eating disorders. AN and BN samples tend to be adolescent females (Garfinkel, Lin, Goering, Spegg, Goldbloom, & Kennedy, 1996b); BED samples are generally older and comprised of a greater percentage of males compared to AN and BN samples (Devlin et al., 2003). Documentation of sample ages and genders allows for generalizability across the eating disorder literature. Clinicians may struggle to incorporate findings from samples that exclude participants normally seen in real-life presentations.

Race and socio-economic status have been linked to distinct pathologies in eating disorder samples (Striegel-Moore, Schreiber, Lo, Crawford, Obarzanek, & Rodin, 2000). Eating disorders have generally been perceived as Caucasian women's diseases, and few studies have specifically reported data in samples that include a substantial subset of minority participants (Rosen, Silberg, & Gross, 1988). However, girls and women of color have higher rates of eating disorders than previously expected (Crago, Shisslak, &

Estes, 1996), and deserve to be represented across the literature. Ethnic groups differ in systematic ways in response to therapy (Chui, Safer, Bryson, Agras, & Wilson, 2007), and scores on eating disorder assessments have been shown to vary by race and SES (Striegel-Moore et al., 2000). Clearly, documentation of racial and socio-economic status is necessary for interpreting and generalizing outcomes presented in research.

Diagnostic subtyping is an important consideration when evaluating the presentations of eating disorder disturbances found in the AN and BN literature. AN subtypes of restricting and binge/purge are markedly different in psychiatric comorbidity (Casper & Davis, 1977; Laessle, Wittchen, Fichter, & Pirke, 1989), and in distress and severity of eating disorder symptoms (Neigo, Pratt, & Agras, 1997). Individuals engaging in bingeing and purging behaviors are more likely to have poorer outcomes and greater impulsivity (Garner, Shafer, & Rosen, 1992; Rossiter & Agras, 1990). BN subtypes differ with regard to type of compensatory behavior present. BN-purging type involves vomiting or laxative abuse, while BN-non-purging type includes excessive calorie intake or exercise. Compensatory behaviors have been highlighted as a primary factor in functional assessment, with particular behaviors associated with more severe outcomes (Spoor et al., 2007; Vitousek, Watson, & Wilson, 1998). Therefore, some assessment of subtypes represented within a sample can provide valuable information with regards to generalizability and comparability across samples.

A history of a previous eating disorder diagnosis is common across eating disorder diagnoses (Keel & Mitchell, 1997). Eating disorder diagnostic duration and migration are predictive of course and outcome in AN and BN populations, and are therefore relevant sample characteristics to eating disorder research consumers. Longer

lifetime duration of BN predict more chronic course and greater general psychopathology (Fahy & Russel, 1993; Keel, Mitchell, Miller, Davis, & Crow, 1999; Reas, Williamson, Martin, & Zucker, 2000). Similarly, an investigation may include individuals who have a lifetime history of another eating disorder, which alters the profile of the sample in significant ways (Godart et al., 2002). Research documentation and assessment of previous eating disorder status in participant samples may help to clarify population profiles represented in the research.

Psychiatric comorbidity has considerable relevance in providing generalization restrictions. Additionally, comorbid Axis I and II disorders present an aggravating factor in the course and treatment of eating disorders (Bulik, Sullivan, Joyce, Carter, & McIntosh, 1998; Herzog, Nussbaum & Marmor, 1996; Saccomani, Savoini, Cirrincione, Vercillino, & Ravera, 1998). Some researchers suggest that additional psychiatric comorbidities exacerbate the general impairment and decrease treatment responsiveness of eating disorder patients (Braun, Sunday, & Halmi, 1994). Axis I and II psychopathology has been associated with higher levels of severity in eating disorder symptoms (Spindler & Milos, 2007). Therefore, documentation of these comorbidities in patient samples is of value to consumers of eating disorder literature, as uncontrolled comorbidities may skew data, muddle meaningful interpretations, and limit an article's interpretive value.

In general, sample characteristics are likely to vary across a literature base. Researchers may strictly implement inclusionary and exclusionary guidelines to achieve distinct samples with clear generalizability, or may consider more heterogeneous samples as a reflection of real-life presentations. Regardless, clear documentation of sample

characteristics must be included in the literature, as these variables have been implicated in the predictive course, outcome, and response to treatment of eating pathology.

Diagnostic Methods

Research is only as generalizable as the sample it selects. Large differences may arise between samples when selection procedures are drastically different across the literature. Researchers evaluating eating disorders may choose from a variety of diagnostic tools and methodologies. Self-report questionnaires, self-monitoring with food diaries, and structured and semi-structured interviews all have strengths and weaknesses for assessment of eating disorder pathology.

Self-report questionnaires. It is difficult to investigate the accuracy of self-reporting of eating disorder symptoms, largely due to the secretive nature of the disorders (Fairburn & Wilson, 1993). However, comparisons of the agreement between diagnostic instruments results in greater symptomology reported in self-report questionnaires than with other methods (Fairburn & Wilson, 1993; Wilfley, Schwartz, Spurrell, & Fairburn, 1997), which may lead researchers to assume a more severe sample is represented.

Food diaries. Self-monitoring with food diaries can provide a detailed assessment of circumstances of eating disorder symptoms. However, these measures rely on the patient for accurate recall and honest reporting of eating behaviors. In addition, these methods allow patients to define criteria by any definition they deem appropriate. For instance, individuals engaging in bulimic binges may choose to define a binge as any amount eaten that may be perceived as excessive, ignoring *DSM* requirements of an

objectively large amount of food given a circumstance. Additionally, several sources of potential inaccuracy with self-monitoring have been identified (Rosen & Srebnik, 1990), and self-monitoring tends not to corroborate well with self-reported recall of past events (Rossiter, Agras, & Telch, 1992). Dietary records have been demonstrated to be unreliable in both normal weight and obese populations (Black, Kehrberg, Flumerfelt, & Schlosser, 1997), and research suggests that the accuracy of retrospective recall significantly diminishes over time (Schoeller, 1990). Finally, the act of recording daily food intake may actually alter intake patterns in significant ways (Dennis, Ernst, Hjortland, Tilloston, & Grambsch, 1980).

Semi-structured and structured interviews. Clinical structured or semi-structured interviews provide the most systematic means of establishing diagnostic criteria (Black & Wilson, 1996; Cooper & Fairburn, 1987; Fairburn & Beglin, 1994). Clear definitions of symptoms and systematic question probing results in high reliability (Carter, Aime, & Mills, 2001; Garner, 2002). However, some common instruments do not assess for duration of key symptoms, despite requirements in criteria set forth in the *DSM* for AN, BN, and BED (the Composite International Diagnostic Interview (CIDI), for example, fails to make provisions for BED thresholds), and criteria definitions may not be systematic in all instruments. As criteria may differ across assessment instruments (Hudson, Hiripi, Pope, & Kessler, 2007), familiarity is required with each structured or semi-structured interview to determine the level of agreement between instruments.

Definition of Diagnostic Criteria

With researchers varying in their adherence to diagnostic criteria definitions, documentation of such adherence is vital in the implication of generalizability of the research sample. Body weight, binge behavior, and purging behavior are implicated as valuable aspects of each of the eating disorder pathologies, and must be assessed across diagnostic categories of AN, BN, and BED. While the *DSM* provides some guidelines and suggestions on how these criteria may be operationalized, it is unclear to what extent these guidelines are followed.

Body weight. Significant weight loss is emphasized in order to receive the diagnosis of AN, and is generally believed to be an indicator of severity of the diagnosis (Herzog, Schellberg, & Deter, 1997). Additionally, weight loss must be assessed in order to distinguish between anorexia nervosa – binge/purge subtype and the diagnosis of BN. The *DSM* suggests a body weight at less than 85% of one's expected weight to meet diagnosis of AN (APA, 2000). However, researcher definition of "significant weight loss" is dependent upon how one defines an individual's "expected weight." The means of defining weight is largely left up to experimenter preference, is often times not cited in the literature, or is cited differently by a number of different sources. Research and clinical variations include references to the Metropolitan Life Insurance Company weight charts (1959), American Insurance Industry's *Build and Blood Pressure Study* (Society of Actuaries, 1959), and Body Mass Index (BMI) measures, all of which provide divergent rationale and have significant generalizability difficulties (Oehlschlagel-Akiyoshi, Malewski, & Mahon, 1999). This variation in the definition and measurement of weight

may create heterogeneity in research samples and unsystematically affect outcomes. Additionally, though an assessment of body weight is not necessary for diagnostic purposes in BED research, body weight is considered to be an indicator of severity of pathology in BED samples (Devlin et al., 2003). There appears to be no single standard for defining body weight across the eating disorder diagnostic categories.

Binge. The occurrence of discrete binge eating episodes exists across the diagnostic spectrum of eating disorders and has become a valuable criterion in the major disorders of AN (binge/purge subtype), BN, and BED. Considerable controversy exists over the most appropriate definition of binge eating (Fairburn & Wilson, 1993), and variability in research is common (Beglin & Fairburn, 1992; Pratt, Niego, & Agras, 1998; Telch, Pratt, & Niego, 1998). The *DSM* defines binge episodes as eating an amount of food that is definitely larger than most individuals would eat under similar circumstances (APA, 1994, 2000). However, researchers may choose to adopt a purely subjective definition, in which a binge is considered to include any amount of food that violates a patient's idea of dietary control or creates anxiety about gaining weight (Schlundt & Johnson, 1990), a purely objective definition, in which caloric intake is measured, or a social-circumstantial approach which takes into account the amount of food consumed according to the circumstance (Garfinkel et al., 1995). BN and BED women are more likely to define binge eating only by a sense of loss of control over eating (Telch et al., 1998), as it appears to be associated with considerable emotional distress (Neigo et al., 1997). Unfortunately, the various methods of binge data collection utilized in BN and BED research have been shown to produce only moderate levels of agreement

(Bartholome, Raymond, Lee, Peterson, & Warren, 2006). Furthermore, research has shown important differences with regard to the definition of a binge as it relates to associated psychopathology (Fairburn & Wilson, 1993; Garner et al., 1992; Rossiter & Agras, 1990), impulsivity, and affective instability (Williamson et al., 2001).

Compensatory behavior. The *DSM* defines compensatory behaviors as the induction of vomiting, misuse of laxatives and diuretics, or excessive exercise (APA, 2000). An individual exhibiting any one of these behaviors may be considered to meet diagnostic criteria. However, research demonstrates that specific compensatory behaviors may not, in fact, be interchangeable, and may result in distinct pathological profiles. Laxative abuse has been associated with lower self-esteem (Kovacs & Palmer, 2004), greater eating pathology (Pryor, Wiederman, & McGilley, 1996), greater general psychopathology such as depression and personality disorders (Bulik, Sullivan, Joyce, & Carter, 1995; Mitchell, Boutacoff, Hatsukami, Pyle, & Eckert, 1986), greater impulsivity (Bruce, Koerner, Steiger, & Young, 2003), and self-harming behaviors (Anderson, Carter, McIntosh, 2002). Additionally, laxative misusers have been found to differ from non-laxative users on a range of sample characteristics including age and duration of the disorder (Fairburn & Cooper, 1982). Excessive exercise is associated with anxious/obsessional temperament and personality characteristics (Shroff et al., 2006), longer inpatient treatment (Solenberger, 2001), and shorter time to relapse (Strober et al., 1997). Researchers have begun to tap into the importance of compensatory behaviors — suggesting that the impact of such behaviors on psychosocial impairment may be greater than that of binge eating (Spoor et al., 2007). Therefore, the presence and type of

compensatory behaviors represented in research samples may drastically alter the outcome data, and researchers should attempt to assess and document the presence of this variable in their participant samples.

Adherence to Criterion Thresholds

Specific criteria and criteria thresholds in eating disorder literature receive varying degrees of acceptability or emphasis among researchers. However, by virtue of their inclusion in the *DSM*, diagnostic criteria sets are expected to be associated with degrees of pathology for eating disordered individuals. Diagnostic threshold criteria for amenorrhea, body weight, frequency of binge episodes, and frequency of compensatory behaviors are often arbitrarily required for participation in research samples. Though the *DSM* provides threshold guidelines, researchers may arbitrarily maintain thresholds at more lenient levels, or may not require the criterion to be met at all.

Amenorrhea. Amenorrhea is considered to be associated with greater eating disorder pathology and lower weight (Copeland, Sacks, & Herzog, 1995), and AN diagnostic status is dependent upon the symptoms of amenorrhea for three consecutive months (APA, 2000). However, some researchers suggest that amenorrhea lacks clinical significance and adherence to this criterion precludes otherwise comparable individuals from participation in research samples (Bunnell, Shenker, Nussbaum, Jacobson, & Cooper 1990; Thaw, Williamson, & Martin, 2001; Williamson, Gleaves, & Savin, 1992). Specifically, amenorrhea has occurred in individuals who have not lost a significant amount of weight, while some individuals continue to menstruate despite extremely low weight (Theander, 1970). Non-amenorrheic individuals who otherwise meet diagnosis for

AN may be comparable to full-AN individuals on a variety of measures of eating pathology (Thaw et al., 2001). However, based on their inability to meet this criterion, these individuals are relegated to the eating disorder not otherwise specified (EDNOS) category. The degree to which amenorrhea is used in inclusionary and exclusionary decisions is not documented and variability in this regard may result in heterogeneous AN (as well as, EDNOS) samples.

Body weight. The *DSM* requires significant underweight status—85% less than an individual’s expected weight—to accompany a diagnosis of AN (APA, 2000). Clinicians have reported significant impairment in functioning in individuals who were unable to meet the weight guideline (Anderson, Bowers, Watson, 2001), leading some to question the criterion’s validity. Investigators have argued that the body weight criterion for a diagnosis of AN is arbitrarily determined and clinically irrelevant (Anderson et al., 2001; Watson & Anderson, 2003). Therefore, significant variability may exist in the utilization of this criterion as a participant requirement, and resulting heterogeneity in the literature is likely.

Binge frequency. Considerable controversy exists over the most appropriate means of measuring binge eating (Fairburn & Wilson, 1993; Spoor et al., 2007). Discrete binge episodes are considered relevant if their occurrence and frequency are established at a clinically meaningful twice-a-week threshold. However, there is substantial variability in the presentation of this symptom, and binges that occur at high rates or last for extended periods of time may be difficult to formally separate into discrete periods. Binge episodes in BED may be particularly difficult to quantify, as they are unlikely to be

terminated by compensatory behaviors (Grilo et al., 2001). Rather than discerning between episodes, some researchers choose to assess binge frequency by number of days on which bingeing occurred, introducing yet another potential measurement difference across the literature.

Additionally, many researchers consider the binge frequency requirement to be too high when determining impairments in functional outcomes (Spoor et al., 2007). Some evidence suggests that individuals who binge less often than twice a week may be similar on most relevant dimensions to the more frequent bingers. BN women failing to meeting binge frequency requirements did not score lower on general measures of psychopathology than those with the threshold diagnosis (Crow, Agras, Halmi, Mitchell, & Kraemer, 2002), and showed no significant differences in psychosocial functioning and health care utilization (Johnson, Spitzer, & Williams, 2001). However, frequency of binge eating episodes has been associated with cluster B personality disorders, as well as Axis I substance-related and anxiety disorders (Spindler & Milos, 2007). Although some researchers have questioned the value in upholding such a controversial threshold requirement, others maintain that key differences exist between threshold and sub-threshold patients. As controversy continues concerning such threshold requirements, the degree to which this criteria threshold is retained across the literature is unknown.

Further, an individual must meet this frequency of bingeing behaviors for a period of three months to warrant a BN diagnosis, and for a period of six months to warrant a diagnosis of BED (APA, 2000). Although some researchers promote a strict adherence to the *DSM* duration requirements of bingeing behavior, others may be more lenient in their requirements. However, longer duration of binge behaviors is significantly related to

chronic course, general psychopathology, and poor outcomes (Fahy & Russel, 1993; Keel et al., 1999; Swain et al., 1991). Therefore, the degree to which this important variable frequency is standardized is of clinical and research relevance.

Compensatory behavior frequency. The assessment of compensatory behaviors is pivotal in the diagnosis of BN and BED, and reaches clinical significance at a twice-a-week threshold. While the diagnosis of BED is based on a disordered eating profile in which bingeing occurs in the absence of compensatory behaviors, many BED individuals exhibit irregular compensatory behaviors at low levels (APA, 2000). Some researchers permit BED research samples to exhibit low levels of compensatory behaviors by delineating a compensatory behavior threshold. Others show no tolerance toward the presence of compensatory behaviors in BED research samples. The presence of these behaviors in research samples should be documented, as compensatory behaviors have been found to be indicative of particular psychopathologies (Garfinkel et al., 1995). More frequent and longer duration of compensatory behaviors suggest a more chronic course with a poor outcome (Fahy & Russell, 1993). BED researchers should be documenting tolerance for irregular compensatory behaviors for their prospective importance to the future understanding of this diagnostic category.

Purpose of Study and Expected Findings

Empirical literature reviews can identify differences in diagnostic practices being documented in published research. Such reviews help to clarify the considerable inconsistencies associated with literature domains (Armstrong, Channell, McGrath, & Maieritsch, 1998). Presently, the eating disorder literature contains numerous discrepant

findings, and outcome studies are often difficult to integrate with other studies in the literature base. Various diagnostic methods, symptom definitions, and criteria requirements are utilized across the eating disorder literature. Though the eating disorders of AN and BN have generated a considerable amount of research, the extent of consistency regarding diagnostic methodologies and documentation across the literature is unknown. Controversies regarding the clinical utility of certain criteria in AN and BN exist and clinical researchers are likely to operationalize criteria differently from one another – we expect variability with regard to *DSM* adherence, especially with the newer BED category. Documentation of any variability in diagnostic or sampling procedures should greatly increase the value of research across both time and inevitable revisions to diagnostic criteria sets.

BED is considered a provisional diagnosis. The validity of this diagnostic status remains in debate, as research is needed to gain further understanding of the presentation. Therefore, clarification regarding sampling procedures across the eating disorders would be useful, and may help to reconcile the current literature state.

The purpose of the present study is to review articles publishing with AN, BN, and BED populations in order to document diagnostic and sampling practices used by researchers in each of these subdomains of the eating disorder literature. Assessment and comparison of the use and documentation of exclusionary and inclusionary criteria, key characteristics of eating disorder samples, symptom definitions, and adherence to *DSM* criteria thresholds across target articles could help researchers reconcile conflicting research findings and may promote a more explicit documentation approach to eating disorder literature.

It was hypothesized that researchers would vary with regard to documentation of diagnostic procedures. Explicit details regarding diagnostic criteria are, most likely, not presented in eating disorder research. Further, variability with regard to particular controversial criteria is likely to be found across diagnoses. Given BED status as a provisional diagnosis, researchers working in this area are likely to provide more explicit detail of sampling procedures and participant characteristics than is provided for more established eating disorder diagnoses.

CHAPTER II

METHODS

Materials

Sets of articles were located via three separate searches of the Web of Science search engine using keywords *Anorexia Nervosa* or *anorexic*, *Bulimia Nervosa* or *bulimic*, and *Binge Eating Disorder*. The Web of Science search engine (<http://scientific.thomson.com/products/wos/>) provides access to current and retrospective information from approximately 8,700 high impact journals. Journals are evaluated by editorial committees comprised of individuals considered expert in their given field, and offered inclusion in the search engine based on a journals ability to meet particular standards suggesting high quality research. Articles are then indexed according to discipline of the journal source, and may be sorted according to a variety of factors, including number of citations. The number of citations represents the number of citations a journal has accumulated in all years on Web of Science, regardless of years in which a search is specified.

Searches were conducted across psychiatric and psychology (subtype 'general') journals published between 2000 and 2007, and the resulting articles were ranked according to the number of times a given article had been cited. The top 100 most cited articles as of December 2007 were obtained for each subdomain. The top thirty most

cited articles meeting the previous guidelines for each of the three disorders were selected for this study. Only articles meeting the following guidelines were included in the study: (1) The article sampled participants formally diagnosed with the eating disorder of interest (Anorexia Nervosa, Bulimia Nervosa, or Binge-Eating Disorder), and (2) the diagnosis was a factor in how the results were presented and discussed. Epidemiological studies were excluded from the study because the diagnosis was a dependent variable in such studies. Five articles that included multiple diagnostic groups and that were among the 30 most cited articles for more than 1 diagnostic group. Thus, 90 research samples/procedures from a total of 85 articles were coded.

Procedure

A coding system was developed to check diagnostic practices documented across articles. Separate coding sheets were used for Anorexia Nervosa (AN – see Appendix A), Bulimia Nervosa (BN – see Appendix B), and Binge-Eating Disorder (BED – see Appendix C). The coding system prompts for demographic information, diagnostic procedures, criteria definition, criteria thresholds and utilization, and comorbid diagnoses.

Reliability was established for the coding system through two phases. The first phase involved reliability training. Research team members (one graduate student in a master's degree program in clinical psychology accredited by the Master's in Psychology Accreditation Council [MPAC], and two undergraduates majoring in psychology) were trained on the coding system by reviewing ten articles per disorder separate from the ones being used in the current study. An expert coder (faculty advisor, licensed psychologist)

provided the benchmark for the training articles. During reliability training, items on the coding sheet were revised if a consensus could not be established regarding coding.

Initial interrater agreements for training materials were computed $[(\# \text{ of agreements}) / (\# \text{ of agreements} + \# \text{ of disagreements})]$. As necessary, additional training articles were coded until interrater agreements reached $>.90$ for the ten most recently coded articles.

Phase two involved the actual coding of the articles. During this phase, 30% of the first author's articles were randomly selected for coding by another team member. Disagreements were resolved through consensus, and at no point did agreements drop below $.90$.

CHAPTER III

RESULTS

Overview of Analyses

Data were entered and analyzed using the Statistical Package for the Social Sciences (SPSS) version 15 on a PC platform. ANOVA procedures were selected for parametric data (e.g., number of participants) and the Cramér's V coefficient (V) was selected for non-parametric (e.g., study type) data. Cramér's V is similar to phi and "can be interpreted as an index that measures the strength of the association between two variables" (Healey, 2002, p. 322). For the purposes of this study, the Cramér's V coefficient was used to measure the association between nominal-level variables. The Cramér's V coefficient is useful in its generalizability across tables of varying sizes (more than two rows and more than two columns) and is not affected by sample size (AcaStat Software, 2003; Healey, 2002). Therefore, significant levels are not artificially inflated by large sample sizes. Researchers previously used these analyses to allow for the discovery of differences and trends in the diagnostic practices reported across the Attention-Deficit Hyperactivity Disorder literature (Dawkins, 2004; Hartley, 2003). Additionally, the Fisher's Exact Test/chi-squared analyses were conducted using GraphPad software (Motulsky, 1995).

Interpreted as a measure of the relative strength of association between at least two variables, the Cramér's *V* coefficient ranges from 0 (no association) to 1.00 (perfect association) (AcaStat Software, 2003). Therefore, the closer the coefficient is to 1.00, the stronger the relationship. A limitation with Cramér's *V* is the difficulty in forming meaningful interpretations of values between 0.00 and 1.00 (Healey, 2002). The values cannot be interpreted as anything other than a relative strength of association.

The Cramér's *V* coefficient was computed to determine the existence of any significant differences on a number of non-parametric variables across the eating disorder diagnoses. First, overall trends were examined for each of the diagnoses. Next, studies were compiled and significant differences were examined between each of the three diagnostic categories.

Fisher's Exact test was utilized to determine differences between diagnoses on a number of variables in which only two of the diagnoses shared the particular feature (e.g., use of subtyping). Fisher's Exact test is a more suitable solution than chi-square analyses, as it does not require reliance on large sample approximations, as such approximations would be invalid for the present study.

ANOVA procedures were utilized to analyze diagnostic category differences for a small number of parametric variables (e.g., number of participants). ANOVA procedures may not ordinarily be used for data analysis of variables such as the mean age of participant samples due to failure of the data to meet typical assumptions. It was used in this study because it is more powerful than its nonparametric counterparts. A more powerful statistic is desirable here because establishing significant differences is of value to readers who wish to ascertain the comparability of research samples across diagnoses.

Because this was an exploratory study, p values between .05 and .10 are reported as indicating a trend towards significance, while p values less than .05 are regarded as significant.

Types of Studies Reviewed

Study Type

Regardless of diagnostic category, the most highly cited eating disorder articles were focused on the further assessment of the eating disorders, rather than specific treatment outcomes. Overall, 75.6% of eating disorder samples were in articles coded as “Assessment or Descriptive” in nature. However, Cramér’s V analyses showed nearly significant differences among the diagnostic categories with regard to the proportion of assessment or descriptive studies and treatment or intervention studies. BED articles were somewhat more likely than AN or BN articles to produce studies in which treatment and intervention were the primary purposes, $V = .256, p = .052$. Table 1 presents the overall percentages of types of studies reviewed for each diagnostic category.

Table 1

Type of Studies Reviewed Overall

	AN (<i>n</i> = 30)	BN (<i>n</i> = 30)	BED (<i>n</i> = 30)
Assessment/ Descriptive	83.3%	83.3%	60.0%
Treatment/ Intervention	16.7%	16.7%	40.0%

Investigatory Department

Primary author affiliations were used to determine the academic department responsible for the research investigation, coded as either ‘psychological,’ ‘psychiatric,’ or ‘other.’ Overall, 68.9% of the research sampled was primarily affiliated with psychiatric departments. Cramér’s V analyses showed no significant differences across diagnostic groups with respect to investigatory department, $V = .225, p = .338$, such that in the case of all diagnoses, psychiatric-affiliated research was most common (see Table 2).

Table 2

Investigatory Department Responsible

	AN (<u>$n = 30$</u>)	BN (<u>$n = 30$</u>)	BED (<u>$n = 30$</u>)
Psychiatric	70.0%	76.7%	60.0%
Psychological	16.7%	13.3%	33.3%
Other	13.3%	10.0%	6.7%

Journal

Nineteen different journals were represented across the top cited articles. No differences were noticed across diagnostic category with respect to journal source, $V = .643, p = .417$. Overall, the top two most cited journals (*American Journal of Psychiatry* and *International Journal of Eating Disorders*) made up 45.5% of the sources. With two highly-cited journals comprising nearly half the sample, the remaining journal sources

were never cited in more than 7% of the total sample. The percentage of representation accumulated by each journal is presented in Table 3.

Table 3

Journal Sources of Top Cited Articles

	AN <u>n = 30</u>	BN <u>n = 30</u>	BED <u>n = 30</u>
American Journal of Psychiatry	26.7%	23.3%	23.3%
International Journal of Eating Disorders	16.7%	20.0%	26.7%
Biological Psychiatry	6.7%	10.0%	6.7%
Archives of General Psychiatry	3.3%	13.3%	6.7%
Psychological Medicine	3.3%	6.7%	3.3%
Journal of Clinical and Consulting Psychology	0.0%	3.3%	10.0%
Journal of Psychiatric Research	0.0%	6.7%	3.3%
Molecular Psychiatry	10.0%	0.0%	0.0%
Journal of Child Psychology and Psychiatry	6.7%	0.0%	3.3%
Journal of Personality Disorders	3.3%	3.3%	3.3%
Journal of Nervous and Mental Disease	3.3%	0.0%	6.7%
European Psychology	3.3%	3.3%	0.0%
Society of Biological Psychiatry	6.7%	0.0%	0.0%
British Journal of Psychiatry	6.7%	0.0%	0.0%
Journal of Abnormal Psychology	0.0%	3.3%	3.3%
Journal of Clinical Psychiatry	0.0%	0.0%	3.3%
Psychosomatic Medicine	0.0%	3.3%	0.0%
Psychological Assessment	3.3%	0.0%	0.0%
Cognitive Therapy and Research	0.0%	3.3%	0.0%

Year

Publication dates across the top cited articles ranged from 2000-2007. The majority of the research samples, 62.2%, were included in articles published in 2000 and 2001, and a downward trend followed, with more recent years accounting for fewer highly cited publications (see Table 4). No significant difference exists across diagnoses with regard to publication date, $V = .229, p = .666$.

Table 4

Publication Years of Top Cited Articles

	<u>AN</u> (<i>n</i> = 30)	<u>BN</u> (<i>n</i> = 30)	<u>BED</u> (<i>n</i> = 30)
2000	33.3%	40.0%	30.0%
2001	30.0%	26.7%	26.7%
2002	20.0%	13.3%	16.7%
2003	16.7%	13.3%	10.0%
2004	0.0%	6.7%	6.7%
2005	0.0%	0.0%	6.7%
2006	0.0%	0.0%	0.0%
2007	0.0%	0.0%	3.3%

Times Cited

The number of citations of the top cited eating disorder articles ranged from 23 to 126 ($M = 48.7$; $SD = 24.1$). No significant differences existed among diagnoses with regard to number of citations, $F(2, 87) = 1.351$, $p = .264$. Table 5 presents the numbers of citations across diagnostic categories.

Table 5

Number of Citations for Top Cited Articles

	AN ($n = 30$)	BN ($n = 30$)	BED ($n = 30$)
<i>M</i>	52.0	51.3	42.9
Median	47.0	42.0	31.0
<i>SD</i>	18.4	26.4	26.2
Minimum Number of Citations	33	29	23
Maximum Number of Citations	112	126	126

Demographics of Eating Disorder Samples

Age

Overall, over half (56.7%) of research samples provided age means. No significant difference was noticed among diagnoses with regard to reporting sample age means, $V = .198$, $p = .171$ (see Table 6). ANOVA procedures showed significant differences among diagnostic categories in the mean age of participants in articles in

which mean age was reported, $F(2, 48) = 22.828, p < .001$. Standard deviations were provided for 53.3% research samples. Significant differences were not found among the diagnoses with regard to tendency to report standard deviations, $V = .197, p = .175$, as well as, mean standard deviations reported by articles $F(2, 45) = 1.477, p = .239$.

Minimum and maximum age of samples was reported for 51.1% and 45.6% of samples, respectively. Articles on BED were significantly more likely than those related to AN or BN to report minimum, $V = .314, p = .012$, and maximum, $V = .395, p = .001$, ages of participants. Proportion of articles reporting age demographic data is presented in Table 6.

Table 6

Reporting of Age Demographics

	<u>AN</u>	<u>BN</u>	<u>BED</u>
<i>M</i>	66.7%	60.0%	43.3%
<i>SD</i>	63.3%	56.7%	40.0%
Minimum Age of Participants	40.0%	40.0%	73.3%
Maximum Age of Participants	30.0%	33.3%	73.3%

When age was reported, significant differences were also found among diagnoses for the mean minimum ages reported for participants, $F(2, 43) = 9.163, p < .001$, as well as for the mean maximum age of participants, $F(2, 38) = 4.191, p = .023$, with

participants in BED groups reportedly having both the highest mean minimum and maximum ages. The mean ages and standard deviations of eating disorder participants as well as the mean minimum and maximum ages are presented in Table 7.

Table 7

Age Demographics Reported

	<u>AN</u>	<u>BN</u>	<u>BED</u>
Mean Age	23.5 (5.1)	25.4 (3.3)	36.3 (8.2)
<i>SD</i>	7.1 (3.5)	6.1 (2.6)	7.9 (2.1)
Minimum Age	14.4 (2.4)	16.1 (2.5)	18.2 (2.6)
Maximum Age	43.9 (16.7)	43.8 (15.1)	55.6 (10.1)

Gender

Gender is a commonly reported participant characteristic in eating disorder literature across diagnostic categories. Gender ratios were reported for 85.6% of eating disorder samples. Cramér’s *V* analyses indicate no significant differences across diagnoses with regard to reporting on gender of participant samples, $V = .089$, $p = .698$. AN researchers reported gender demographics in 90% of articles, while both BN and BED researchers reported gender demographics in 83.3% of articles.

When gender information was provided, females were specifically included in 100% of articles. Cramér’s *V* analyses suggest no differences across diagnoses with regard to inclusion of female participants, $V = .089$, $p = .698$. Overall, 64.5% of articles specifically included a female-only population. Regardless of diagnostic category,

researchers sampled relatively heavily from female populations. Using the ANOVA procedure, no significant differences for mean number of female participants were found across diagnoses, $F(2, 74) = .845, p = .434$.

Eating disorder samples are likely to draw female populations, and males were included in samples at low rates across diagnoses. Males were specifically included in 21.1% of eating disorder samples. None of the sample articles included a male-only eating disorder population. In fact, males were specifically not included in 65.4% of eating disorder samples. However, BED and AN samples were significantly more likely than BN samples to include male participants in research samples, $V = .308, p = .014$. In fact, only one BN sample included male participants. A significant difference across mean number of male participants was found, $F(2, 74) = 4.116, p = .020$, such that BED samples included a greater overall proportion of male participants than either AN or BN samples. The mean numbers of male and female participants are presented in Table 8.

Table 8

Gender Demographics When Gender was Reported

	<u>Females</u>		
	<u>AN</u>	<u>BN</u>	<u>BED</u>
Mean	100.70	75.80	80.44
<i>N</i>	27	25	25
SD	80.173	68.955	71.312
	<u>Males</u>		
	<u>AN</u>	<u>BN</u>	<u>BED</u>
Mean	1.59	0.16	6.84
<i>N</i>	27	25	25
SD	3.016	.800	14.910

When males were specifically included, mean numbers of male participants ranged from 4.00 in BN and 4.78 in AN samples, to 19.00 in BED samples. No significant differences among samples with regard to mean number of male participants when males were specifically included, $F(2, 16) = 2.311, p = .131$.

Race

Overall, race information was provided in 31.1% of all eating disorder samples. Race information for AN and BN samples were reported 23.3% and 16.7% of the time, respectively. Race information for BED samples was reported 53.3% of the time.

Cramér's V analyses showed significant differences among diagnoses, $V = .344, p = .005$, such that BED samples were more likely than AN or BN samples to report race information for participants. Table 9 presents proportion of samples providing race information.

Twenty-nine percent of samples reported using non-Caucasian participants. African Americans were included in 13.3% of the participant samples across diagnoses. Cramér's V analyses indicated that BED samples were significantly more likely to include African Americans, $V = .424, p < .001$, and Native Americans, $V = .263, p = .045$, than the other two diagnostic categories. However, significant differences were not found in the proportion of samples within each diagnostic category that used Hispanic Americans (included in 4.4% of all studies) or Asian Americans (included in 4.4% of all studies).

ANOVA procedures indicate that no significant differences between mean number of African Americans, $F(2, 22) = .133$, Hispanic Americans, $F(2, 21) = .121$, Native Americans, $F(2, 21) = .390$, and Asian Americans, $F(2, 21) = .055$, were provided according to diagnostic category. Regardless of diagnostic category, minority races were unlikely to represent a large proportion of research samples. The mean numbers of non-Caucasian participants are presented in Table 9.

Table 9
Race Demographics Reported

<u>African Americans (n = 25)</u>			
	<u>AN</u>	<u>BN</u>	<u>BED</u>
<i>M</i>	0	5.33	20.53
<i>n</i>	7	3	15
<i>SD</i>	0	5.033	27.961
<u>Hispanic Americans (n = 24)</u>			
	<u>AN</u>	<u>BN</u>	<u>BED</u>
<i>M</i>	.29	2.33	.756
<i>n</i>	7	3	14
<i>SD</i>	7.56	3.215	1.336
<u>Native Americans (n = 24)</u>			
	<u>AN</u>	<u>BN</u>	<u>BED</u>
<i>M</i>	0.0	0.0	.20
<i>n</i>	7	2	15
<i>SD</i>	0.0	0.0	.414
<u>Asian Americans (n = 24)</u>			
	<u>AN</u>	<u>BN</u>	<u>BED</u>
<i>M</i>	9.00	37.67	0.0
<i>n</i>	7	3	14
<i>SD</i>	23.374	62.660	0.0

Socio-Economic Status

Overall, 73.3% of eating disorder samples failed to assess socio-economic status (SES) in research populations. SES was reported for 33.3% of AN samples, 16.7% of BN samples, and 30.0% of BED samples. Cramér's V analyses found no significant difference in assessment of SES among diagnostic categories, $V = .163$, $p = .303$. Regardless of diagnostic category, SES was relatively rarely assessed for participant samples.

Referral Source

Overall, a referral source for participants was indicated in 51.1% of eating disorder samples. Cramér's V analyses revealed no differences with regard to reporting recruitment sample across diagnoses, $V = .190$, $p = .186$. Researchers specifically reported selecting from inpatient, outpatient, mixed (both inpatient and outpatient) or community samples. Researchers specifically recruited from clinical (inpatient or outpatient) populations in 51.1% of research samples. BED samples were slightly less likely to be recruited from clinical populations. No significant differences among diagnoses were found with regard to sampling from clinical populations, $V = .157$, $p = .329$ (see Table 10).

Outpatient populations were the most commonly recruited referral source, accounting for 44.4% of sample source across the diagnoses. Researchers did not differ with regard to recruitment of outpatient samples across diagnostic categories, $V = .084$, $p = .730$. Clinical inpatient participants were included in eating disorder samples in 20.0% of articles. Cramér's V showed significant differences among diagnostic categories with

regard to the proportion of inpatient populations sampled, $V = .312, p = .013$.

Specifically, BED samples were significantly less likely to draw from inpatient populations compared with AN or BN samples.

Community samples were specifically recruited in 25.6% of the overall sample articles. BED researchers were more likely to draw from community samples than from clinical populations. Cramér's V showed significant differences among diagnostic categories with regard to recruitment of community samples, $V = .355, p = .003$. BED samples were significantly more likely than the other diagnostic categories to recruit from community populations (see Table 10).

Table 10

Referral Populations Reported

	<u>Referral Populations</u>		
	AN (<i>n</i> = 30)	BN (<i>n</i> = 30)	BED (<i>n</i> = 30)
In Treatment	56.7%	56.7%	40.0%
Community	10.0%	20.0%	46.7%

Subtypes

Subtypes were specifically assessed in 61.7% of AN and BN samples. Fischer's Exact test indicates a significant difference between AN and BN samples, $p = 0.0326$, such that articles using AN samples were more likely than articles using BN samples to report an assessment of subtypes across participants. In 76.7% of AN samples, subtypes were specifically assessed whereas subtypes were assessed in only 46.7% of BN samples.

Overall, 33.3% of AN and BN samples, both subtypes were specifically included in the participant sample. Both subtypes were included in 50% of AN samples, and in only 16.7% of BN samples. AN samples are significantly more likely than BN samples to include both subtypes, $p = .006$ (Fisher's Exact test).

Anorexia Nervosa subtypes. One or more subtypes were diagnosed in 76.7% of AN samples. AN-binge/purge subtype was specifically included in samples in 66.7% of all AN articles, and in 87.0% of all AN articles which assessed for subtype status. AN-restricting subtype was specifically included in samples in 60.0% of all AN articles, and in 78.2% of all articles which assessed for subtype status. No significant difference was found between proportion of AN articles assessing for each subtype ($p = .0838$, Fisher's Exact test).

Bulimia Nervosa subtypes. One or more BN subtype was diagnosed in 46.7% of BN articles. BN-purging subtype was specifically included in samples in 46.7% of all BN articles, and in 100% of articles assessing subtype status. BN-non-purge subtype was specifically included in samples in 16.7% of all BN articles, and in 35.7% of articles which assessed for subtype status. Fisher's Exact test revealed significant difference between BN subtypes, $p = .0125$, such that BN-non-purge subtype was not likely to be represented in BN research samples.

Eating Disorder Duration

Overall, eating disorder duration was assessed in 47.8% of eating disorder samples. According to Cramér's V analyses, differences among diagnostic category

approached significance, $V = .252, p = .058$. BED samples were somewhat less likely to report an assessment of duration of eating disturbances than either AN or BN samples. Articles with AN and BN samples reported disorder duration assessment 56.7% of the time; articles with BED samples only reported disorder duration assessment 30% of the time.

History of Other Eating Disorder

Overall, an assessment of history of other eating disorder was reported for 28.9% of eating disorder samples. Diagnostic categories did not differ with regards to assessment of a history of other eating disorder, $V = .243, p = .071$ (see Table 11). Cramér's V analyses suggested no differences among diagnostic categories on inclusionary, $V = .170, p = .271$, or exclusionary, $V = .181, p = .277$, use of this variable. AN, BN, and BED research similarly assessed for and used a history of a previous eating disorder in sampling decisions. Regardless of diagnostic category, articles were more likely to report the inclusion of participants with a history of a previous eating disorder diagnosis than they were to exclude such participants. Table 11 shows the frequency of reported assessment of history of previous eating disorder diagnoses across diagnostic categories and proportion of articles utilizing this variable in inclusionary or exclusionary decisions.

Table 11

Treatment of History of Other Eating Disorder

	AN (<i>n</i> = 30)	BN (<i>n</i> = 30)	BED (<i>n</i> = 30)
Assessed	36.7%	36.7%	13.3%
Included	26.7%	30.0%	13.3%
Excluded	10.0%	6.7%	0.0%
Not Assessed	63.3%	63.3%	86.7%

Dual Diagnoses

Overall, the existence of dual diagnoses for participants with eating disorders was not mentioned or referred to in any way (either as inclusionary or exclusionary criteria) in 38.9% of eating disorder samples. Cramér's *V* analyses did not show any significant differences among the diagnoses in the proportion of studies that mentioned comorbidity as a factor in sampling procedures, $V = .179, p = .235$.

Comorbid diagnoses were specifically allowed in 37.8% of samples. According to Cramér's *V* analyses, no significant differences were indicated among the diagnoses in the proportion of samples that specified whether dual diagnoses were permitted or allowed, $V = .125, p = .842$. Likewise, no significant differences were noticed across diagnoses with regard to inclusion of specific Axis I or Axis II disorders. Table 12 presents the percentages of studies that permitted specific comorbid diagnoses.

Table 12

Dual Diagnoses Allowed in Eating Disorder Populations

	<u>AN</u>	<u>BN</u>	<u>BED</u>	<u>V</u>	<i>p</i>
Substance-Related Disorders	26.7%	6.7%	23.3%	.224	<i>ns</i>
Mood / Depressive Disorders	23.3%	10.0%	33.3%	.230	<i>ns</i>
Anxiety Disorders	16.7%	10.0%	23.3%	.146	<i>ns</i>
Personality Disorders	16.7%	10.0%	20.0%	.115	<i>ns</i>

Exclusionary Criteria

Overall, exclusionary criteria were mentioned in 40.0% of eating disorder samples, with no differences existing among diagnoses in tendency to mention exclusionary criteria, $V = .056$, $p = .870$. Percentages of excluded disorders ranged from 5.5% excluding anxiety disorders, 8.9% excluding depressive disorders, 17.8% excluding substance-related disorders and personality disorders, to 32.2% excluding psychotic disorders. Pregnancy was specifically excluded in 25% of BN and BED samples, and medical problems that influence eating or weight were specifically excluded in 14.5% of samples. No significant differences exist among diagnoses with regard to any specific exclusionary criteria or diagnoses. The percentages of samples within each diagnostic category that excluded specific diagnoses or criteria are presented in Table 13.

Table 13

Exclusionary Diagnoses Mentioned

	<u>AN</u>	<u>BN</u>	<u>BED</u>	<u>V</u>	<u>p</u>
Substance-Related Disorders	13.3%	13.3%	26.7%	.164	<i>ns</i>
Mood / Depressive Disorders	10.0%	10.0%	6.7%	.055	<i>ns</i>
Anxiety Disorders	10.0%	3.3%	3.3%	.137	<i>ns</i>
Personality Disorders	20.0%	13.3%	20.0%	.082	<i>ns</i>
Psychotic Disorders	30.0%	23.3%	43.3%	.178	<i>ns</i>
Pregnancy	0.0%	20.0%	30.0%	.381	<i>ns</i>
Medical Problem Influencing Eating	6.7%	13.3%	23.3%	.195	<i>ns</i>

Diagnostic Methods Used

Diagnostic Methods Reported

Diagnostic methods were documented for 70.0% of eating disorder samples, and were significantly more likely to be reported for BED samples, $V = .309$, $p = .014$. For both AN and BN categories, the proportion of articles reporting diagnostic methods was 60.0%; however, 90.0% of BED articles reported diagnostic methods.

Diagnoses Confirmed by Researchers

Authors from the majority of articles (57.8%) reported confirming the eating disorder diagnoses of the participant samples. BED articles were significantly more likely

than AN or BN articles to clearly indicate confirmation of diagnosis, $V = .393, p = .008$.

Table 14 presents the proportion of articles reporting a confirmation of diagnosis by researchers within each diagnostic group.

Table 14

Diagnosis Confirmed By Researchers

	AN (<i>n</i> = 30)	BN (<i>n</i> = 30)	BED (<i>n</i> = 30)
Confirmed	50%	43.3%	80.0%
Not Confirmed	43.3%	36.7%	6.7%
Unclear	6.7%	20.0%	13.3%

Self-Report Questionnaires

Questionnaires were specifically used for diagnostic purposes in 7.8% of the eating disorder samples. Researchers reporting on AN, BN, and BED subjects utilized questionnaires for diagnosis in 0.0%, 3.3%, and 20.0% of articles, respectively. Overall, when questionnaires were used, 57.1% preferred the Eating Disorder Examination-Questionnaire (EDE-Q), 28.6% used the Binge Eating Scale (BES), and 14.3% used the Questionnaire on Eating and Weight Patterns (QEWP). Significant differences with regard to use of questionnaire were found among diagnostic categories, such that BED samples were more likely to be diagnosed by questionnaires, $V = .327, p = .008$.

Food Diaries

Food diaries were specifically used for diagnostic purposes in 4.4% of the sample articles. Researchers reporting on AN, BN, and BED subjects utilized food diaries for diagnosis in 0.0%, 3.3%, and 10.0% of articles, respectively. No differences among diagnostic categories were found regarding use of food diaries, $V = .202$, $p = .160$.

Interviews

Structured interviews. Structured interviews were specifically used for diagnostic purposes in 64.4% of the eating disorder samples. The Eating Disorder Examination (EDE) and Structured Clinical Interview for DSM (SCID) made up 64.9% of the diagnostic interviews utilized. The use of structured interviews approached significance according to diagnostic category, $V = .230$, $p = .093$. BED articles was somewhat more likely to report use of structured interviews than AN or BN articles. Researchers diagnosis AN and BN samples utilized structured interviews in 56.7% of the articles; however, BED samples were diagnosed with structured interviews 80.0% of the time. Regardless of diagnosis, structured interviews were the most commonly utilized diagnostic tool.

Semi-structured interviews. Semi-structured interviews were specifically used for diagnostic purposes in 16.7% of samples. A trend toward significance was found regarding use of semi-structured interviews across diagnostic categories, $V = .247$, $p = .064$. BED researchers were somewhat more likely to use semi-structured interviews (13.3%) than either AN (3.3%) or BN (0.0%) researchers. BED researchers exclusively

utilized the EDE, whereas AN researchers used the Structured Interview for Anorexia and Bulimia Nervosa.

Nine different structured or semi-structured interviews were employed in the top cited eating disorder articles. Differences approached significance regarding specific interview tool utilized across diagnoses, $V = .651$, $p = .087$. The proportion of each interview used in research samples in which interviews were the diagnostic instrument are presented in Table 15.

Table 15

Interview Used

	<u>AN</u> <i>(n = 16)</i>	<u>BN</u> <i>(n = 17)</i>	<u>BED</u> <i>(n = 24)</i>
Structured Clinical Interview for DSM	25.0%	17.6%	54.2%
Eating Disorder Examination	12.5%	35.3%	37.5%
Schedule for Affective Disorders and Schizophrenia	25.0%	11.8%	0.0%
Composite International Diagnostic Interview	12.5%	5.9%	4.2%
Diagnostic Interview Schedule	12.5%	5.9%	0.0%
Structured Interview for Anorexia and Bulimia Nervosa	6.25%	11.8%	0.0%
EATATE Interview	6.25%	5.9%	0.0%
Interview for Diagnostic Eating Disorders –IV	0.0%	0.0%	4.2%
Eating Disorder Family History Interview	0.0%	5.9%	0.0%

Definition of Diagnostic Criteria

Definition of Body Weight

Body weight was specifically assessed in 65.6% of all research samples. Though a specific body weight criterion is required for only AN diagnoses, only 40.0% of AN samples were specifically assessed for body weight. The majority of BN samples (66.7%) and BED samples (90.0%) were specifically assessed for body weight. BED articles were significantly more likely to specifically report an assessment of body weight than the other diagnostic categories, $V = .430, p < .001$.

The majority of research samples (52.2%) specifically assessing body weight did so using a Body Mass Index (BMI) calculation. Body weight was assessed using BMI calculation for 58.3% of AN samples, 80.0% of BN samples, and 88.9% of BED samples. Cramér's V analyses showed significant differences among diagnoses regarding use of BMI as a measurement of body weight, $V = .463, p < .001$, such that BN and BED reported use of BMI more so than AN. Metropolitan Life Insurance Company tables were used as a body weight index in 12.2% of the samples, with no differences existing among AN, BN, and BED samples, $V = .048, p = .902$ (33.3%, 20.0%, and 11.1%, respectively).

Definition of Binge

Seventy-nine of the overall 90 samples specifically included diagnostic categories requiring the assessment of binge behavior for diagnosis. These samples included AN- binge/purge subtype, BN, and BED participant populations. The majority of these 79 samples (77.2%) provided diagnosis by a specific tool (e.g., a structured or semi-structured interview, or questionnaire) and provided no further embellishment of

diagnostic definition for binges. Thus, criteria definitions for bingeing were significantly more likely to be inferred through the use of a diagnostic tool, than to be specifically defined by researchers, $V = .268, p = .023$ (see Table 16).

In some cases, criteria definitions were not defined at all. Research populations in which binges were not defined at all made up 22.8% of the sample. BED articles were significantly more likely than the other two diagnostic categories to include some reference to binge definitions used in diagnostic procedures, $V = .307, p = .024$, either specifically provided by researchers or inferred through the use of a particular diagnostic tool.

Binge definitions were specifically provided for diagnosis in 17.7% of the 79 samples. Significant differences were found among diagnostic categories, $V = .272, p = .054$, such that AN – binge/purge articles were less likely than BN or BED articles to specifically provide the definition used to assess presence of bingeing (see Table 16).

Table 16

Reporting Binge Definition

	ANBP (<i>n</i> = 19)	BN (<i>n</i> = 30)	BED (<i>n</i> = 30)
Specifically provided by researchers	0.0	20.0%	26.7%
Inferred from diagnostic tool	63.1%	50.0%	66.7%
Not mentioned at all	36.9%	10.0%	6.7%

Overall, binges were defined objectively for 3.8% of the samples, with no differences existing among diagnoses with respect to the use of an objective definition, $V = .253, p = .082$. Out of the 79 selected articles, subjectively defined binges were never specifically permitted in any diagnostic procedure. The diagnostic use of a social-circumstantial binge definition was provided in 16.4% of samples, and a trend towards significant was recognized, $V = .253, p = .080$, such that BN and BED researchers were somewhat more likely to define binges social-circumstantially. Binges were specifically required to be defined by a sense of lack of control in 11% of all research samples, and lack of control was significantly more likely to be incorporated into a binge definition in BED samples, $V = .305, p = .025$. Researchers including the AN diagnostic subtype of binge/purge never provided any binge definition. Proportion of articles utilizing each binge definition is presented in Table 17.

Table 17

Binge Definitions Reported

	ANBP (<i>n</i> = 19)	BN (<i>n</i> = 30)	BED (<i>n</i> = 30)
Objectively	0.0%	0.0%	3.3%
Subjectively	0.0%	0.0%	0.0%
Social- Circumstantially	0.0%	20.0%	23.3%
Lack of Control	0.0%	6.7%	23.3%

Definition of Compensatory Behavior

Forty-nine of the overall 90 samples specifically included diagnostic categories requiring the assessment of compensatory behaviors. These samples included AN – binge/purge subtype and BN populations. The majority of these 49 samples (73.5%) failed to provide an explicit assessment and definition of compensatory behaviors within the sample. In such cases, 55.6% of the researchers assessed for diagnosis by a specific tool (e.g., an interview or questionnaire), and no further explanation of compensatory behavior assessment was provided. The other 17.9% of cases offered no comment regarding diagnostic methods in general, or compensatory behaviors in particular. No differences regarding the diagnostic reporting of compensatory behavior assessment and definition were found among diagnoses, $p = .205$ (Fisher’s Exact test). Table 18 presents the proportion of samples addressing compensatory behavior definitions needed for ANBP and BN diagnosis.

Table 18

Reporting Diagnostic Definition of Compensatory Behavior

	<u>Compensatory Behavior Definition Provided for Diagnosis</u>		
	<u>ANBP</u> <u>(n = 19)</u>	<u>BN</u> <u>(n = 30)</u>	<u>Overall</u> <u>(n = 49)</u>
Specifically provided by researchers	15.8%	33.3%	26.5%
Inferred from diagnostic tool	47.4%	36.7%	55.6%
Not mentioned at all	36.8%	30.0%	17.9%

Additionally, ten BED samples specifically reported assessment of the presence of irregular compensatory behaviors within research populations. In these BED articles, compensatory behaviors were not required for diagnostic purposes, but were mentioned rather as a sample characteristic. Therefore, 59 total samples (ANBP, BN, and BED) required some assessment of compensatory behaviors. Of BED samples, 70.0% provided a specific definition of compensatory behavior present. Fisher's Exact test suggest BED researchers were significantly more likely than AN researchers to provide specific compensatory behavior definitions for samples, $p = .0108$. Therefore, despite diagnostic relevance to AN researchers including ANBP subtype, BED researchers were significantly more likely to include specific compensatory behavior definitions in research samples. BED articles, however, did not differ significantly from BN articles, $p = .0663$, with regard to likelihood to provide specific compensatory behavior definitions.

Purging was the most commonly named compensatory behavior in research samples. Of the 59 samples including the assessment of compensatory behaviors, purging, laxative abuse, intake restriction, and excessive exercise were specifically named in 30.5%, 13.6%, 10.2%, and 10.2%, respectively. BED and BN articles more likely to specifically name purging as the compensatory behavior present in research samples, $V = .320$, $p = .049$. When compensatory behaviors were specifically defined, purging was named 90% of the time, where as laxative abuse, intake restriction and excessive exercise were included in the definition 40%, 30%, and 30% of the time, respectively.

BED articles were significantly more likely than AN and BN articles to report laxative abuse, $V = .613$, $p < .001$, intake restriction, $V = .597$, $p < .001$, and excessive

exercise, $V = .597$, $p < .001$, in research samples that exhibited compensatory behaviors.

Table 19 presents the proportion of articles defining specific compensatory behaviors in samples.

Table 19

Compensatory Behavior Definitions Reported

	<u>Compensatory Behavior Definitions</u>		
	<u>ANBP</u> <u>(n = 19)</u>	<u>BN</u> <u>(n = 30)</u>	<u>BED</u> <u>(n = 10)</u>
Purging	15.8%	30.0%	60.0%
Laxative Abuse	5.3%	3.3%	60.0%
Intake Restriction	0.0%	3.3%	50.0%
Excessive Exercise	0.0%	3.3%	50.0%

Adherence to DSM Criteria

Adherence to Amenorrhea Criterion

Some variability existed regarding researcher adherence to the amenorrhea criterion for AN diagnosis. Amenorrhea was specifically not required in 20.0% of AN articles. More commonly, diagnosis was reached through a specific diagnostic tool, and amenorrhea was not specifically mentioned (43.3%). However, in 36.7% of AN articles, no mention of diagnostic methods or amenorrhea was included at all.

Adherence to Body Weight Criterion

Some variability existed regarding researcher adherence to the body weight criterion for AN diagnosis. Body weight was specifically required in 16.7% of articles, and specifically not required in 20.0% of articles. The majority of AN articles did not specifically mention the body weight criterion at all (56.7%). However, in 23.3% of AN articles, no mention of diagnostic methods or body weight was mentioned at all.

Adherence to Binge Frequency Criterion

Seventy-nine samples specifically included diagnostic categories requiring the assessment of binge behavior for diagnosis, including the assessment of binge frequency. Overall, the majority of samples (79.7%) were diagnosed by a specific tool (e.g., a structured or semi-structured interview, or questionnaire) and researchers provided no further embellishment of required diagnostic thresholds for binge behavior. However, specific binge frequency requirements were mentioned in 25.3% of all samples. BED and BN articles were significantly more likely than AN articles to specifically provide binge frequency requirements for inclusion in research samples, $V = 334, p = .012$. Proportions of articles providing binge frequency requirements are presented in Table 20.

Table 20

Reporting Binge Frequency Requirements

	<u>Binge Frequency Mentioned</u>		
	<u>ANBP</u> <u>(n = 19)</u>	<u>BN</u> <u>(n = 30)</u>	<u>BED</u> <u>(n = 30)</u>
Mentioned by researcher	0.0%	33.3%	36.6%
Inferred through diagnostic tool	40.0%	40.0%	60.0%
Not mentioned at all	23.3%	26.7%	3.3%

Overall, binge frequency was defined by *days* on which bingeing occurred for 10.1% of samples in which bingeing frequency was assessed. Binge frequency was defined by day in 0.0% of ANBP diagnoses, in 11.1% of BN diagnoses, and in 63.6% of BED diagnoses. BED articles were more likely than either AN or BN articles to provide binge frequency in terms of *day*, $V = .345$, $p = .009$. Binge frequency was defined by *episode* of binges in 16.5% of articles in which binge frequency was assessed. Binge frequency was defined by episode in 0.0% of ANBP diagnoses, in 88.9% of BN diagnoses, and in 36.4% of BED diagnoses. BN was significantly more likely than AN or BED to provide binge frequencies by *episodes* in which bingeing occurred, $V = .340$, $p = .006$. Thus, when binge frequency requirements were reported, BN articles showed a definite trend towards determining frequency by episode, whereas BED utilized both days and episodes as a measure of binge frequency.

Adherence to Compensatory Behavior Frequency Criterion

Forty-nine samples specifically included diagnostic categories requiring the assessment of compensatory behaviors for diagnosis, including the assessment of compensatory behavior frequency. The majority of all samples (55.1%) provided diagnosis by a specific tool (e.g., a structured or semi-structured interview, or questionnaire) and provided no further embellishment of required diagnostic thresholds for compensatory behavior. However, for 18.4% of all samples, researchers specifically mentioned a compensatory behavior frequency requirement. Fisher’s Exact test revealed no difference between AN and BN articles for the proportion specifically providing a compensatory behavior frequency requirement, $p = .069$. Table 21 presents the proportion of articles providing compensatory behavior frequency for diagnosis.

Table 21

Reporting Compensatory Behavior Frequency Requirements

	<u>ANBP</u> <u>(n = 19)</u>	<u>BN</u> <u>(n = 30)</u>
Mentioned by researcher	10.5%	16.7%
Inferred through diagnostic tool	57.9%	53.3%
Not mentioned at all	31.6%	30.0%

Ten of the BED articles allowed for irregular presentation of compensatory behaviors in research samples. Therefore, 59 total samples addressed compensatory behaviors. Twenty percent of BED articles which allowed irregular presentation of

compensatory behaviors specifically reported a compensatory behavior frequency requirement. No significant differences was noticed among ANBP, BN, and BED articles with regard to tendency to report requirements of compensatory behavior frequency, $V = .096, p = .760$.

Some variability existed across diagnosis with regard to parameters specified for compensatory behavior frequency. 100% of AN samples and 60% of BN samples were required to meet a compensatory behavior frequency of 2x/week for a duration of 3 months. BN researchers also employed more lenient (1x/week for a duration of 3 months) and more stringent (2x/week for a duration of 6 months) requirements. BED researchers allowed for irregular compensatory behaviors to be exhibited in frequencies of less than 2x/month, as well as less than 5x/6 months. However, the majority of BED researchers (80%) specifically allowing for irregular compensatory behaviors failed to delineate compensatory behavior frequency cut-off requirements.

Table 22

Compensatory Behavior Frequency Cut-Offs Mentioned

	ANBP <i>(n = 2)</i>	BN <i>(n = 5)</i>	BED <i>(n = 2)</i>
2x/week for 3 months	100.0%	60.0%	-
1x/week for 3 months	0.0%	20.0%	-
2x/week for 6 months	0.0%	20.0%	-
Less than 2x/month	-	-	50.0%
Less than 5x/6 months	-	-	50.0%

CHAPTER IV

DISCUSSION

The purpose of the study was to provide an empirical analysis of diagnostic and sampling practices in a set of the most cited journal articles with anorexia nervosa, bulimia nervosa, and binge-eating disorder samples. As expected, both consistencies and inconsistencies in reporting diagnostic procedures were found across the three diagnostic categories. Overall, BED studies were more likely to clearly report diagnostic methods, criteria threshold requirements, and criteria definitions than AN or BN studies. However, regardless of diagnostic category, most articles neglected to provide important types of information regarding participant selection procedures.

The discussion below focuses on participant selection procedures and the representativeness of report samples, definition of and adherence to *DSM* criteria. Finally, limitations of the current study are provided and recommendations are made regarding documentation and sampling practices in order to increase the utility of the eating disorder research literature.

Types of Studies Reviewed

The top cited articles received between 23 and 126 citations. Nearly half of all top cited articles came from two journals (*American Journal of Psychiatry* and *International Journal of Eating Disorders*). As would be expected, the majority of highly-cited articles

have early publication dates, and citation numbers decrease in more recently published articles. Sixty percent of the samples were published in 2000 and 2001, and have likely accumulated citations over time.

Articles were primarily assessment-oriented and descriptive in nature, and were generally produced by a psychiatry-affiliated first author. Assessment research is expected in a literature base that is filled with various diagnostic controversies and inconsistencies, as researchers struggle to find the most appropriate means of quantifying the pathological presentations represented in the eating disorders. However, researchers publishing within the BED diagnostic category were more likely to generate studies in which treatment and intervention were the primary purposes.

Demographics Reported Across Eating Disorder Samples

Clinicians and clinical researchers value the presentation of detailed demographic information in research articles. Demographic information regarding age and gender of participants was consistently reported in the eating disorder samples. A combination of age characteristics (mean, standard deviation, minimum and maximum) was reported in approximately half of the eating disorder samples. All diagnostic categories were equally likely to report age mean and standard deviation. However, BED articles were more likely than other diagnoses to report minimum and maximum age ranges for participant samples. Unfortunately, regardless of diagnostic category, race and socio-economic status (SES) were rarely mentioned for participant samples, even though they are considered important variables for several eating disorder factors (Striegel-Moore et al., 2000).

When demographic information was included, AN and BN samples were significantly younger than BED populations, and had both younger minimum and maximum ages. Thus, BED research samples were drawn from a significantly older population than either of the other disorders, representing the common spread across the three disorders. In general, anorexic and bulimic samples are generally young Caucasian females; while BED samples are older, and more likely to include a greater proportion of male participants, and a great proportion of non-Caucasians. Because BED affects a greater proportion of males in the general population (Devlin et al., 2003), this difference is expected and suggests greater utility of the developing literature base for BED.

To some degree, these analyses suggest the top cited articles provide accurate characterizations of the common, real-life clinical presentations of the diagnostic categories with respect to demographics. However, a number of deficiencies were found within and among diagnostic categories with respect to demographics. BN articles almost never included male participants, and non-Caucasians were severely under-represented in the top cited AN and BN samples. Consumers of the research literature are likely having their conceptual base most strongly and somewhat narrowly influenced by articles not containing less common or ignored characteristics.

Though racial information was provided by few studies, it was more likely to be provided for BED participants. BED articles were most likely to report including African American and Native American participants. In general, African Americans were the most commonly included non-Caucasian race in participant samples across diagnostic categories. Hispanic-Americans and Asian-Americans participants were reported in virtually none of the eating disorder articles. Further, non-Caucasians made up a small

proportion of participants when included in samples. While racial distinctions provide important implications regarding BED presentation, the same is true of anorexic and bulimic participants. Though non-Caucasian presentations of eating disorders were once thought to be rare, research has indicated the current trend is an increase in prevalence across other racial groups (Crago et al., 1996). Additionally, Caucasian women experience greater eating and body image concerns than ethnic minority women, and different racial groups respond differently to intervention approaches (Wildes & Emery, 2001). Unfortunately, AN and BN articles largely ignore the documentation of race, or sample exclusively from Caucasian populations. Documentation of these kinds of characteristics in the research literature could potentially help clinicians make better decisions with their particular clients.

Other important characteristics of research samples were also rarely documented. It is well-documented that the majority of the reviewed eating disorder literature uses samples from outpatient or inpatient hospitalized populations. However, relatively few affected individuals in the community actually seek treatment. While, studies investigating specialized clinical populations have advanced the scientific understanding of eating disorders, their findings may not be readily generalized to other populations in which eating disorders are less severe (Grilo, Devlin, Cachelin, & Yanovski, 1997). These findings support previous critiques that eating disorder researchers ignore important referral considerations and sample almost exclusively from more severely afflicted populations. Inpatient participants were more likely to be included in AN and BN samples, while outpatient participants were more likely to be included in BED samples.

Additionally, current subtyping schemas suggest that subtypes within a diagnostic category are significantly different on key variables. AN researchers commonly assessed for and included subtype information, but BN researchers were much less likely to do so. Further, AN subtypes were evenly represented in samples, while BN researchers were skewed in reporting on the purging subtype significantly more often than the non-purging subtype. Though the eating disorder subtypes differ qualitatively from one another, AN subtypes present more distinct phenomenological differences that may be easier for researchers to diagnostically categorize. Specifically, the AN subtypes can be assigned based on whether any compensatory behavior is present. However, the BN subtypes require further assessment for the type of compensatory behavior (i.e., purging vs. restricting). In BN, research supports the distinction of these compensatory behaviors because they carry the most impact with regard to functional outcomes. The finding that many BN researchers do not provide specific information on the types of compensatory behavior observed in their participants greatly reduces the generalizability of their results. Additionally, some epidemiological studies imply that the BN- non-purge subtype is, in fact, not a rare phenomenon at all – and may actually be more frequent than the purging type in general populations (Garfinkel et al., 1996a). The current analyses suggest this subtype is largely, and needlessly, ignored across BN samples.

Other sample characteristic variables often unreported include duration of eating disorder symptoms, history of other eating disorders, and the presence of comorbid Axis I and II disorders. Articles with BED samples were less likely to provide eating disorder duration than AN and BN articles. Previous research suggests that duration of eating disorders predict functional impairment and treatment outcomes in AN and BN samples.

Diagnostic categories were also unlikely to provide information regarding participants' previous history with eating disturbances and current comorbidity. Researchers have recognized the considerable diagnostic migration found between AN and BN diagnoses is associated with greater distress and greater general psychopathology. Regardless of diagnostic category, approximately 70% of samples were not assessed for previous eating disorder diagnosis. However, regardless of diagnostic category, articles assessing for a previous history of an eating disorder were more likely to report the inclusion of participants with a history of a previous eating disorder diagnosis than they were to report exclusion of such participants.

Most eating disorder articles mentioned comorbidity in some manner (either as inclusionary or exclusionary criteria) and rarely were researchers specifically interested in studying participants with only an eating disorder diagnosis. Nearly forty percent of samples specifically included comorbid diagnoses. Most commonly, regardless of eating disorder category, participants had comorbid substance-related, mood, or anxiety disorders. These diagnoses are often found in real-life eating disorder samples, suggesting that research samples are representative of real-life presentations in this respect. Exclusionary diagnoses were mentioned in less than half (40.0%) of all eating disorder samples. Most commonly, researchers excluded individuals with psychotic disorders.

Establishment of Eating Disorder Diagnosis

Regardless of diagnostic category, the majority of eating disorder articles provided some documentation of diagnostic procedure. Eating disorder articles are likely

to report diagnostic procedures, to confirm diagnosis of samples, and to do so through the use of structured interviews. BED researchers are most likely to provide detailed documentation of diagnostic procedures, and are most likely to confirm the diagnostic status of samples. Because BED is a provisional diagnosis, research diagnostic detail is extremely powerful in ensuring long-term value to research contributions.

Regardless of diagnostic category, researchers relied heavily on structured interviews to establish eating disorder diagnoses. Most commonly, the use of the Structured Clinical Interview for DSM (SCID) and the Eating Disorder Examination (EDE) were cited. The EDE is considered the gold-standard for eating disorder diagnosis, and adherences specifically to *DSM* criteria sets (Garner, 2002). Semi-structured interviews were cited less often as diagnostic tools. Only two semi-structured interviews were used – BED articles utilized the Eating Disorder Examination in semi-structured form, AN articles utilized the Structured Interview for Anorexia and Bulimia Nervosa (SIAB). Questionnaires were rarely used as a means of establishing eating disorder diagnoses; however, BED articles were more likely to employ such diagnostic methods. Eating Disorder Examination – Questionnaire (EDE-Q) was the most highly utilized questionnaire. However, considerable discrepancies of symptom report on EDE-Q and EDE have been documented (Carter et al., 2001). Researchers did not rely heavily on food diaries to establish diagnoses. In general, reliance on structured interviews should increase homogeneity of research samples in the literature and improve comparability of obtained results.

Definition of Diagnostic Criteria

Diagnostic definitions are most likely dealt with indirectly in the reviewed eating disorder literature. Researchers generally prefer mention of a diagnostic tool, with the assumption that valuable diagnostic information, such as criteria definitions, may be inferred. However, in the research samples for top cited articles, 12 different standardized diagnostic tools were mentioned with varying degrees of frequency. Assumed equivalence across many measures may be unwarranted. Additionally, familiarity with diagnostic tools is necessary to effectively evaluate how important diagnostic considerations were dealt with in samples. Thus, despite overt criteria controversy and variously employed diagnostic definitions, criteria definitions were more likely to be inferred through the use of a diagnostic tool than to be specifically defined by researchers.

Body weight assessment is required by *DSM-IV-TR* (APA, 2000) for AN and BN diagnoses, but just how to conduct the assessment is left up to researcher preference. Body weight was reported by many BED researchers, even though this is not required in the current criteria set. In fact, BED articles were more likely than AN or BN articles to provide assessment and report of body weight in samples. Though the majority of samples were assessed for body weight by means of the BMI calculation, AN articles were less likely than others to do so.

Across all eating disorder categories, body weight assessment was provided most commonly by Body Mass Index (BMI) calculation, or through use of the Metropolitan Life Insurance Company tables. These two tools can result in different classification of

individuals and this can affect the generalizability of obtained results (Oehlschlagel-Akiyoshi et al., 1999).

Few articles provided a specific binge definition for diagnosis. When a definition was reported, binges were defined by most AN, BN, and BED researchers using the *DSM* sanctioned definition, requiring an social-circumstantial. BED articles were most likely to specifically incorporate another important *DSM* requirement—the sense of ‘lack of control’—into binge definitions. Articles never allowed for a purely subjective definition of bingeing to meet diagnostic criteria, and rarely expected objective definitions to fully capture bingeing in diagnostic samples. Surprisingly, AN articles specifically including the binge/purge subtype never provided any binge definition in diagnostic procedures.

Specific compensatory behaviors have been shown to be indicative of distinct pathologies in eating disorders (Bulik et al., 1995; Kovacs & Palmer, 2004; Mitchell et al., 1986; Pryor et al., 1996). However, compensatory behaviors were unlikely to be specifically defined for eating disorder samples in this study. BED samples specifically allowing for irregular compensatory behaviors were more likely than other diagnostic categories to define specific compensatory behaviors. Despite diagnostic relevance for AN and BN samples, BED articles provided greater detail regarding type of compensatory behaviors represented.

When compensatory behaviors were specifically detailed in samples, purging was most common, and other compensatory behaviors were infrequently included. However, most commonly, researchers reached a diagnosis using specified diagnostic tools, but otherwise did not mention types of compensatory behavior present in their sample. In such cases, compensatory behavior assessment must be inferred through familiarity with

the diagnostic tool. Most diagnostic tools are likely to probe for any number of compensatory behaviors that may be present in an individual. The majority of researchers (approximately 70%) did not further clarify the specific compensatory types represented in samples. Therefore, citation of a specific diagnostic tool allows for no reliable assumptions regarding types of compensatory behaviors that may present in research samples.

A sizeable portion of sample articles failed to provide any mention of diagnostic assessment generally, or compensatory behaviors in particular. Therefore, in general, eating disorder articles in which compensatory behaviors are a required diagnostic criterion are unlikely to provide valuable information regarding the types of compensatory behaviors represented within samples.

Criterion Adherence

Some variability exists regarding researcher emphasis on particular criterion within eating disorder diagnostic sets. Despite the diagnostic controversies associated with eating disorder criteria, the majority of researchers fail to specify how they handle potentially controversial criterion thresholds. Heterogeneity in the handling of these thresholds severely reduces the enduring value of research, as criteria sets shift to accommodate evolving understandings of eating disorder diagnostic categories. Researchers generally do mention diagnostic tools, perhaps with the assumption that valuable details regarding diagnostic decisions may be inferred from knowing which interview or questionnaire was used.

AN researchers are faced with controversy surrounding amenorrhea and body size diagnostic criteria. Unfortunately, AN researchers typically did not specifically mention amenorrhea or body weight requirements in diagnostic procedures. The present analyses suggest that variability is common with regard to treatment of these criteria – whether researchers specifically include, specifically exclude, or refrain from mentioning these criteria in their samples. When specifically mentioned, researchers vary considerably regarding use of these controversial criteria, and in 20% of articles, researchers specifically did not require participants to meet amenorrhea or body weight criteria despite *DSM* criteria sets (APA, 2000).

Binge frequencies are specifically delineated for AN – binge/purge, BN, and BED diagnostic sets in the *DSM*. However, eating disorder articles are unlikely to specifically mention binge frequency requirements. This is especially troublesome for AN, for which no *DSM* requirements are established. Therefore, AN – binge/purge subtype samples are including individuals who exhibit bingeing behavior in an undocumented range of frequencies. This practice introduces a large amount of heterogeneity into the literature if not corrected in future research.

Compensatory behavior frequencies are delineated only for BN diagnostic sets. However, AN – binge/purge samples display compensatory behaviors, and *DSM* currently requires no frequency requirement to diagnose the subtype. When compensatory behaviors were assessed, relatively few samples (approximately 15%) reported frequency parameters. Compensatory behavior frequency was unlikely to be mentioned, regardless of diagnostic category. When mentioned, some variability existed with regard to the frequency cut-off employed in studies. Both AN articles and the

majority of BN articles mentioning a compensatory behavior frequency requirement followed the suggested 2x/week for a period of three months that is typical of BN diagnostic requirements for the same behaviors. However, some BN researchers chose to be more or less lenient with their requirements.

BED samples may exhibit irregular compensatory behaviors. Researchers allowing for compensatory behaviors at ‘irregular’ frequencies failed to mention their particular frequency requirement in 80% of BED articles. This is especially problematic, as no current standard is in place for allowing such behaviors in BED populations. If researchers are lenient towards compensatory behaviors, the clarity of the distinction between BN and BED samples may become muddled. Clear documentation of this variable is necessary in maintaining a cohesive understanding of BED research.

Summary of AN, BN, and BED Diagnostic Procedures

The top cited AN literature is assessment-oriented or descriptive. Anorexic samples represented in the top cited literature are likely to document some diagnostic methodology for establishment of a diagnosis. However, AN researchers allow considerable variability within samples across several dimensions. Researchers are systematically silent with regard to criterion definitions. Body weight assessment is unlikely to be mentioned, and is assessed with different methods. Binge and compensatory behaviors are left undefined. Criterion adherence is unsystematic with regards to amenorrhea, body weight, and binge/purge frequencies. Results from the present analyses suggest AN literature is likely to assess for subtype status, and is likely to include both subtypes in research samples. However, researchers including the AN –

binge/purge subtype never provided any information regarding the diagnoses of the subtype. Some delineation of appropriate frequency and intensity cut-off of bingeing and compensatory behaviors is needed, as researchers may be including individuals with any number of frequencies and intensity of the target behaviors important in this subtype.

The top cited BN literature is assessment-oriented or descriptive-driven. Bulimic samples represented in the top cited literature are like to be providing some documentation of diagnostic methods. However, BN researchers often left important variables unmentioned. Results from present analyses suggest BN literature is unlikely to assess for subtype status, and significantly favors one subtype (BNP) over another in research samples. BN researchers utilized standard *DSM* binge definitions, and preferred to measure binge behavior in terms of *days* in which it occurred rather than *episodes*. However, variability may be included within samples with regard to use of compensatory behavior definitions and thresholds, which were only rarely mentioned. When mentioned, researchers varied with regard to criterion thresholds, despite clear *DSM* provisions.

The top cited BED literature is more likely than others to be treatment-oriented. Binge-eating disorder samples represented in the top cited literature are more likely than other diagnostic categories to provide detailed descriptions of samples, diagnostic criterion, and criterion thresholds. Further, regardless of diagnostic necessity, BED samples were described on a number of variables including body weight and irregular compensatory behaviors exhibited. When diagnostically relevant criteria were specifically mentioned, BED researchers adhered to *DSM* suggestions regarding binge definition. However, some variability within BED samples was noticed with regard to

measurement of binge behaviors. BED researchers varied with regards to measuring binge frequency by *days* and by *episodes*.

The clinical presentation of BED is often accompanied by irregular compensatory behaviors. In general, BED researchers either failed to mention whether compensatory behaviors were allowed, or specifically excluded compensatory behaviors in samples. However, some researchers specifically allowed compensatory behaviors at irregular frequencies, and BED articles were more likely than other diagnostic categories to specifically define the behaviors seen within research samples. Unfortunately, no standardization was seen with regard to measurement of “irregular” compensatory behaviors frequencies across such samples.

Finally, BED articles utilized questionnaires within diagnostic procedures more so than AN or BN articles. Self-reporting on questionnaires has been shown to provide significantly different presentations of eating disturbances, such that individuals are likely to report greater symptomology in self-report than in standardized interviews.

Limitations of the Current Research

The current research has some significant limitations to consider. The intention of using the “most cited” articles was to examine diagnostic practices and sampling characteristics in influential articles. However, an article may be cited frequently by other researchers but rarely read by practitioners or even valued by readers. Also, an article can become more heavily cited based on factors other than the quality of the reported research. For example, an article may be cited frequently because it has a particular kind of important flaw that other authors want to discuss as a flaw they avoided. Further, it is

not uncommon for researchers to self-cite, or for journals to encourage citations within their journal (Gorman, 2005). Such habits may artificially inflate citation numbers that may not be indicative of the most influential research articles within a given research area. Additionally, an article left out of this sample could be highly influential on researcher or clinician behavior but rarely or never cited.

Methodologically, there were a large number of statistical analyses run on a relatively small amount of data. In such cases, any one individual response may gain an inappropriate amount of power to skew interpretations. Not too much should be made about any individual finding given the rather generous parameters used in detecting potential differences.

This study also relied on results produced through a search on the Web of Science. Web of Science generates references based on a number of factors. Articles are then sorted into categories following an evaluation of content by an editorial review board. From these categories, the present search was conducted. Therefore, the study relies on the search engine's editorial evaluation to have sorted articles into appropriate categories. Highly-cited articles not categorized as "psychological" or "psychiatric," may still have a great deal of influence over the eating disorder literature, but were not included in the present study.

Lastly, observed differences in what ends up in print may reflect not just researcher behavior but editor influences. Researchers may have included detail in original manuscripts that editors request be deleted, for example, out of an editor's desire to conserve space.

Recommendations for Documentation of Diagnostic Practices

Given the controversial landscape of eating disorder literature, clear reports of diagnostic practices by researchers allows for greater evaluation and comparison of research findings. While especially important for emerging diagnostic categories such as BED, it is also important that researchers using established diagnostic categories like AN and BN provide sufficient detail to allow their work to be understood now and to remain valuable as their categories evolve. A minimum standard for research should be to utilize a systematic method for assessing and reporting diagnostic categories. When diverse cases are mixed within a sample, the obtained outcome will be restricted due to within group variability (Kazdin, 1995). Researchers not providing sufficient detail to adequately evaluate sample characteristics do a disservice to research consumers.

Demographic information regarding samples provides important generalization limitations. Consistent and accurate report of sample characteristics, including variables found to be associated with distinctive presentations, should be documented. Males and non-Caucasians are underrepresented in current literature; however, eating disorders can still be functionally impairing for these populations. Research populations across a literature base should, at a minimum, represent the natural spread of important demographics within the diagnoses, as to capture the most common presentations of each disorder.

Eating disorders commonly occur comorbidly with other psychiatric disorders, and “pure” eating disorders are not generally represented in eating disorder literature. However, researchers are not consistently reporting assessment of psychiatric comorbidity, or which dual diagnoses are being including in samples.

When establishing an eating disorder diagnosis, researchers must be clear regarding how participants are being diagnosed and what methods are being used. Though many researchers are providing a diagnostic tool, most criteria remain unaddressed. This is understandable in research, as diagnostic procedures are unlikely to be detailed unnecessarily; however, many criteria are currently under a great deal of scrutiny. The eating disorder literature clearly contains heterogeneity with regard to diagnostic definitions and diagnostic adherence. Researchers should expand upon definitions utilized in the diagnostic procedures; particularly with regard to unknown or lesser known diagnostic tools. Further, if researchers choose not to strictly follow the criteria set forth in the *DSM*, they should thoroughly document any deviations. Such documentation allows for other researchers and clinicians to be aware of the specific standards used in the study, and to accurately evaluate the findings. For example, deviations from amenorrhea, body weight, and behavioral frequencies and durations should always be documented.

Researchers must effectively document criteria definitions and adherence to (or deviation from) *DSM* diagnostic criteria in order to establish comparability of research samples across studies and reduce overall unexplained heterogeneity in the literature. Improved documentation of diagnostic practices in eating disorder literature is crucial and will increase generalizability of research findings and allow for improved clinical utility of experimental outcomes. Clinicians will be better able to utilize a literature that clearly identifies sample characteristics and sample selection procedures.

REFERENCES

- AcaStat Software (2003). *Coefficients for measuring association*. Retrieved May 29, 2003, from <http://www.acastat.com/Handbook/27.html>
- American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed). Washington, DC: Author.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed). Washington, DC: Author.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders*. (4th ed., text revision). Washington, DC: Author.
- Anderson, A. E., Bowers, W. A. & Watson, T. A. (2001). A slimming program for eating disorders not otherwise specified: Reconceptualizing a confusing, residual, diagnostic category. *Psychiatric Clinics of North American*, 24, 271.
- Anderson, C. B., Carter, F. A., & McIntosh, V. V. (2002). Self-harm and suicide attempts in individuals with bulimia nervosa. *Eating Disorders: The Journal of Treatment & Prevention*, 10(3), 227-243.
- Armstrong, K. J., Channell, M., McGrath, A., & Maieritsch, S. (1998, November). *What ADHD diagnostic criteria are reported by JACP authors?* Poster presented at the annual meeting of the Association for the Advancement of Behavior Therapy, Washington, DC.
- Bartholome, L. T., Raymond, N. C., Lee, S. S., Peterson, C. B., & Warren, C. S. (2006). Detailed analysis of binges in obese women with binge eating disorder: Comparisons using multiple methods of data collection. *International Journal of Eating Disorders*, 39, 685-693.
- Beglin, S. J., & Fairburn, C. G. (1992). Evaluation of new instrument for the detection of eating disorders in community samples. *Psychiatry Research*, 44, 191-201.
- Black, D. W., Kehrberg, L. L., Flumerfelt, D. L., & Schlosser, S. S. (1997). Characteristics of 36 subjects reporting compulsive sexual behavior. *American Journal of Psychiatry*, 154, 243-249.

- Black, C. M., & Wilson, T. G. (1996). Assessment of eating disorders: Interview versus questionnaire. *International Journal of Eating Disorders*, 20(1), 43-50.
- Braun, D. L., Sunday, S. R., & Halmi, K. A. (1994). Psychiatric comorbidity in patients with eating disorders. *Psychological Medicine*, 24, 859-867.
- Bruce, K. R., Koerner, M. N., Steiger, H., & Young, S. N. (2003). Laxative misuse and behavioral disinhibition in bulimia nervosa. *International Journal of Eating Disorders*, 33, 92-97.
- Bulik, C. M., Sullivan, P. F., Joyce, P. R., & Carter, F. A. (1995). Temperament, character, and personality disorder in bulimia nervosa. *Journal of Nervous and Mental Disorders*, 183, 593-598.
- Bulik, C.M., Sullivan, P.F., Joyce, P.R., Carter, F.A. & McIntosh, V. V. (1998). Predictors of 1-year treatment outcome in bulimia nervosa. *Comprehensive Psychiatry*, 39, 209-214.
- Bunnell, D. W., Shenker, I. R., Nussbaum, M. P., Jacobson, M. S., Cooper, P. (1990). Subclinical versus formal eating disorders: Differentiating psychological features. *International Journal of Eating Disorders*, 9(3), 357-362.
- Carter, J. C., Aime, A. A., & Mills, J. S. (2001). Assessment of bulimia nervosa: A comparison of interview and self-report questionnaire methods. *International Journal of Eating Disorders*, 30, 187-192.
- Casper, R. C., & Davis, J. M. (1977). On the course of anorexia nervosa. *American Journal of Psychiatry*, 134, 974-978.
- Chui, W., Safer, D. L., Bryson, S. W., Agras, W. S., & Wilson, G. T. (2007). A comparison of ethnic groups in the treatment of bulimia nervosa. *Eating Behaviors*, 8(4), 458-491.
- Cooper, Z., & Fairburn, C. (1987). The eating disorder examination: A semi-structured interview for the assessment of the specific psychopathology of eating disorders. *International Journal of Eating Disorders*, 6(1), 1-8.
- Copeland, P. M., Sacks, N. R., & Herzog, D. B. (1995). Longitudinal follow-up of amenorrhea in eating disorders. *Psychosomatic Medicine*, 57, 121-126.
- Crago, M., Shisslak, C. M., & Estes, L.S. (1996). Eating disturbances among Hispanic and Native American youth. *International Journal of Eating Disorders*, 10, 179-186.

- Crow, S. J., Agras, W. S., Halmi, K., Mitchell, J. E., & Kraemer, H. C. (2002). Full syndrome versus subthreshold anorexia nervosa, bulimia nervosa, and binge eating disorder: A multicenter study. *International Journal of Eating Disorders*, 32, 309-318.
- Dawkins, L. (2004). *An empirical review of researchers' AD/HD diagnostic practices reported in high impact pediatric, psychology, and psychopharmacological journals*. Unpublished master's thesis, Mississippi State University, Starkville.
- Devlin, M. J., Goldfein, J. A., & Dobrow, I. (2003). What is this thing called BED? Current status of binge eating disorder nosology. *International Journal of Eating Disorders*, 34(Suppl), 2-18.
- Dennis, B., Ernst, N., Hjortland, M., Tilloston, J., & Grambsch, V. (1980). The NHLBI nutritional data system. *Journal of American Diet Association*, 77, 641-647.
- de Zwaan, M., Mitchell, J. E., & Seim, H. C. (1994). Eating related and general psychopathology in obese females with binge eating disorder. *International Journal of Eating Disorders*, 15(1), 43-52.
- Eddy, K. T., Keel, P. K., & Dorer, D. J. (2002). Longitudinal comparison of anorexia nervosa subtypes. *International Journal of Eating Disorders*, 31(2), 191-201.
- Fahy, T. A., & Russel, G. F. (1993). Outcome and prognostic variables in bulimia nervosa. *International Journal of Eating Disorders*, 14(3), 135-145.
- Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders*, 16(4), 363-370.
- Fairburn, C. G., & Cooper, J. P. (1982). Self-induced vomiting and bulimia nervosa: An undetected problem. *British Medical Journal*, 284, 1153-1155.
- Fairburn, C. G., Welch, S. L., Doll, H. A., Davies, B. A., & O'Connor. (1997). Risk factors for bulimia nervosa: A community-based case-control study. *Archives of General Psychiatry*, 54, 509-517.
- Fairburn, C. G., & Wilson, G. T. (1993). Binge eating: Definition and classification. In C.G. Fairburn & G.T. Wilson (Eds.), *Binge eating: Nature, assessment, and treatment* (pp. 3-14) New York: Guilford Press.
- Garfinkel, P. E., Kennedy, D. H., & Kaplan, A. S. (1995). Views of classification and diagnosis of eating disorders. *Canadian Journal of Psychiatry*, 40, 445-456.

- Garfinkel, P. E., Lin, E., Goering, P., Spegg, C., Goldbloom, D.S., Kennedy, S., et al. (1996a). Purging and non-purging forms of bulimia nervosa in a community sample. *International Journal of Eating Disorders*, *20*, 231-238.
- Garfinkel, P. E., Lin, E., Goering, P., Spegg, C., Goldbloom, D., Kennedy, S., et al. (1996b). Should amenorrhea be necessary for the diagnosis of anorexia nervosa? Evidence for a Canadian community sample. *British Journal of Psychiatry*, *168*, 500.
- Garner, D. M. (2002). Measurement of eating disorder psychopathology. In C.G. Fairburn & K.D. Brownell (Eds.), *Eating disorders and obesity: A comprehensive handbook* (2nd ed.) (pp. 141-146). Guilford Press: New York.
- Garner, D. M., Shafer, C. L., & Rosen, L. W. (1992). Critical appraisal of the DSM-II-R diagnostic criteria for eating disorders. In S.R. Hooper, G.W. Hynd, & R.E. Mattison (Eds.), *Child psychopathology: Diagnostic criteria and clinical assessment* (pp. 261-303). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.
- Godart, N. T., Flament, M. F., Perdereau, F., & Jeammet, P. (2002). Comorbidity between eating disorders and anxiety disorders: A review. *International Journal of Eating Disorders*, *32*, 253-270.
- Gorman, G. E. (2005). How do we count our chickens? Or do citation counts count? *Online Information Review*, *29*, 581-584.
- Grilo, C. M., Devlin, M. J., Cachelin, F. M., & Yanovski, S. Z. (1997). Report of the National Institutes of Health (NIH) workshop on the Development of Research Priorities in Eating Disorders. *Psychopharmacology Bulletin*, *33*, 321-333.
- Grilo, C. M., Masheb, R. M., & Wilson, G. T. (2001). Subtyping binge eating disorder. *Journal of Consulting and Clinical Psychology*, *69*, 1066-1072.
- Guss, J. L., Kissileff, H. R., Devlin, M. J., Zimmerli, E., & Walsh, B. T. (2002). Binge size increases with body mass index in women with binge-eating disorder. *Obesity Res*, *10*, 1021-1029.
- Hartley, J. M. (2003). *Review of documentation of ADHD diagnostic practices reported in a sample of high impact journals*. Unpublished master's thesis, Mississippi State University, Starkville.
- Healey, J. F. (2002). Association between variables measured at the nominal level. *Statistics: A tool for social researcher* (6th ed.) (pp. 320-330). Belmont, CA: Wadsworth.

- Herpertz-Dahlmann, B., Muller, B., Herpertz, S., & Heussen, N. (2001). Prospective 10-year follow-up in adolescent anorexia nervosa—course, outcome, psychiatric comorbidity, and psychosocial adaptation. *Journal of Child Psychology and Psychiatry*, *42*, 603-612.
- Herzog, D. B., Nussbaum, K. M., & Marmor, A.K. (1996). Comorbidity and outcome in eating disorders. *Psychiatr Clin North Am*, *19*, 843-859.
- Herzog, W., Schellberg, D., & Deter, H. (1997). First recovery in anorexia nervosa patients in the long-term course: A discrete-time survival analysis. *Journal of Consulting and Clinical Psychology*, *65*, 169-177.
- Hsu, L. K., (1980) Outcome of anorexia nervosa: A review of the literature (1954 to 1978). *Archives of General Psychiatry*, *37*, 1041-1046.
- Hudson, J. I., Hiripi, E., Pope, H. G., & Kessler, R. C. (2007). The prevalence and correlates of eating disorders in the national comorbidity survey replication. *Biological Psychiatry*, *61*, 348-358.
- Joiner, T. E., Vohs, K. D., & Heatheron, T. F. (2000). Three studies on the factorial distinctiveness of binge eating and the bulimic symptoms among nonclinical men and women. *International Journal of Eating Disorders*, *27*, 198-205.
- Johnson, J. G., Spitzer, R. L., & Williams, J. B. W. (2001). Health problems, impairment, and illnesses associated with bulimia nervosa and binge eating disorder among primary care and obstetric gynecology patients. *Psychological Medicine*, *31*, 1455-1466.
- Kazdin, A. E. (1995). Scope of child and adolescent psychotherapy research: Limited sampling of dysfunctions, treatments, and client characteristics. *Journal of Clinical Child Psychology*, *24*, 125-140.
- Keel, P. K. & Mitchell, J. E. (1997). Outcome in bulimia nervosa. *American Journal of Psychiatry*, *25*, 768-780.
- Keel, P. K., Mitchell, J. E., Miller, K. B., Davis, T. L., & Crow, S. J. (1999). Long-term outcome of bulimia nervosa. *Archives of General Psychiatry*, *56*, 63-69.
- Klump, K. L., Bulik, C. M., & Pollice, C. (2000). Temperament and character in women with anorexia nervosa. *Journal of Mental and Nervous Disease*, *188*(9), 559-567.
- Kovacs, D., & Palmer, R. L. (2004). The associations between laxative abuse and other symptoms among adults with anorexia nervosa. *International Journal of Eating Disorders*, *36*, 224-228.

- Laessle, R. G., Wittchen, H. U., Fichter, M. M., & Pirke, K. M. (1989). The significance of subgroups of bulimia and anorexia nervosa: Lifetime frequency of psychiatric disorders. *International Journal of Eating Disorders*, 8, 569-574.
- Masheb, R. M., & Grilo, C. M. (2000). Binge eating disorder: A need for additional diagnostic criteria. *Comprehensive Psychiatry*, 41, 159-162.
- Metropolitan Life Insurance Company (1959). New weight standards for men and women. *Statistical Bulletin of the Metropolitan Life Insurance Company*, 40, 1-4.
- Mitchell, J. E., Boutacoff, L. I., Hatsukami, D., Pyle, R. L., & Eckert, E. D. (1986). Laxative abuse as a variant of bulimia. *Journal of Nervous and Mental Disorders*, 20, 13-18.
- Motulsky, H. (1995). *GraphPad: Data analysis and biostatistics software and resources*. Retrieved May 28, 2008: www.graphpad.com
- Myers, L. (2002). *A survey of published researchers: Which ADHD participant characteristics should be documented?* Unpublished master's thesis, Mississippi State University, Starkville.
- Neigo, S. H., Pratt, E. M., & Agras, W. S. (1997). Subjective or objective binge: Is the distinction valid? *International Journal of Eating Disorders*, 22, 291-298.
- Oehlschlagel-Akiyoshi, J., Malewski, P., & Mahon, J. (1999). How to define anorectic weight? *European Eating Disorders Review*, 7, 321-333.
- Pratt, E. M., Niego, S. H., & Agras, W. S. (1998). Does the size of the binge matter? *International Journal of Eating Disorders*, 24, 307-312.
- Pryor, T., Wiederman, M. W., & McGilley, B. (1996). Laxative abuse among women with eating disorders: An indication of psychopathology? *International Journal of Eating Disorders*, 20, 13-18.
- Reas, D. L., Williamson, D. A., Martin, C. K., & Zucker, N. L. (2000). Duration of illness predicts outcome for bulimia nervosa: A long-term follow-up study. *International Journal of Eating Disorders*, 27(4), 428-434.
- Rosen, J. C., Silberg, N. T., & Gross, J. (1988). Eating Attitudes Test and Eating Disorders Inventory: Norms for adolescent girls and boys. *Journal of Consulting and Clinical Psychology*, 56(2), 305-508.
- Rosen, J. C., & Srebnik, D. (1990). The assessment of eating disorders. In P. McRenolds, J.C. Rosen, & G.J. Chelune. (Eds.), *Advances in psychological assessment, Vol 7* (pp. 229-259). New York, NY: Plenum Press.

- Rossiter, E. M., & Agras, W. S. (1990). An empirical test of the DSM-III-R definition of binge. *International Journal of Eating Disorders*, 9, 513-518.
- Rossiter, E. M., Agras, W. S., & Telch, C. F. (1992). The eating patterns of non-purging bulimic subjects. *International Journal of Eating Disorders*, 11, 111-120.
- Saccomani, L., Savoini, M., Cirricione, M., Vercellino, F., & Ravera, G. (1998). Long-term outcome of children and adolescents with anorexia nervosa: Study of comorbidity. *Journal of Psychosomatic Research*, 44, 565-571.
- Schlundt, D. G., & Johnson, W. G. (1990). *Eating disorders: Assessment and treatment*. (pp. 513) Needham Heights, MA: Allyn & Bacon.
- Schoeller, D. A. (1990). How accurate is self-reported dietary energy intake? *Nutritional Review*, 48, 373-379.
- Shroff, H., Reba, L., Thornton, L. M., Klump, K. L., Berrettini, W. H., Brandt, H. et al. (2006). Features associated with excessive exercise in women with eating disorders. *International Journal of Eating Disorders*, 39(6), 454-461.
- Society of Actuaries (1959). *Build and Blood Pressure Study*, Vol. 1. Chicago: Society of Actuaries.
- Solenberger, S. E. (2001). Exercise and eating disorders: A 3-year inpatient hospital record analysis. *Eating Behaviors*, 2(2), 151-168.
- Spindler, A., Milos, G. (2007). Links between eating disorder symptom severity and psychiatric comorbidity. *Eating Behaviors*, 8, 364-373.
- Spoor, S. T. P., Stice, E., Burton, E., & Bohon, C. (2007). Relations of bulimic symptom frequency and intensity to psychosocial impairment and health care utilization: Results from a community-recruited sample. *International Journal of Eating Disorders*, 40, 505-514.
- Striegel-Moore, R. H., Schreiber, G. B., Lo, A., Crawford, P., Obarzanek, E., & Rodin, J. (2000). Eating disorder symptoms in a cohort of 11 to 16-year-old black and white girls: The NHLBI Growth and Health Study. *International Journal of Eating Disorders*, 27, 49-66.
- Striegel-Moore, R. H., Wilson, G. T., & Wilfley, D. E. (1998). Binge eating in an obese community sample. *International Journal of Eating Disorders*, 23, 27-37.

- Strober, M., Freeman, R., & Morrell, W. (1997). The long-term course of severe anorexia nervosa in adolescents: Survival analysis of recovery, relapse, and outcome predictors over 10-15 years in a prospective study. *International Journal of Eating Disorders*, 22, 339-360.
- Stunkard, A. J., & Allison, K. C. (2003). Binge eating disorder: Disorder or marker? *International Journal of Eating Disorders*, 34(Suppl.), 107-116.
- Sullivan, P. F., Bulik, C. M., Fear, J. L., & Pickering, A. (1998). Outcome in anorexia nervosa: A case-control study. *American Journal of Psychiatry*, 155, 939-946.
- Sullivan, P. F., Bulik, C. M., & Kendler, K. S. (1998). The epidemiology and classification of bulimia nervosa. *Psychological Medicine*, 28(3), 599-610.
- Swain, B., Shisslak, C. M., & Crago, M. (1991). Clinical features of eating disorders and individual psychological functioning. *Journal of Clinical Psychology*, 47, 702-707.
- Telch, C. F., Pratt, E. M., & Niego, S. H. (1998). Obese women with binge eating disorder define the term binge. *International Journal of Eating Disorders*, 24, 313-317.
- Thaw, J. M., Williamson, D. A., & Martin, C. K. (2001). Impact of altering DSM-IV criteria for anorexia and bulimia nervosa on the base rates of eating disorder diagnoses. *Eating Weight Disorders*, 6, 121-129.
- Theander, S. (1970). Anorexia nervosa: A psychiatric investigation of 94 female patients. *ACTA Psychiatrica Scandinavica*, 214, 1-194.
- Vitousek, K., Watson, S., & Wilson, G. T. (1998). Enhancing motivation for change in treatment-resistant eating disorders. *Clinical Psychology Review*, 18, 391-420.
- Watson, T. L., & Anderson, A. E. (2003). A critical examination of the amenorrhea and weight criteria for diagnosing anorexia nervosa. *ACTA Psychiatrica Scandinavica*, 108, 175-182.
- Welch, S. L. & Fairburn, C. G. (1994). Sexual abuse and bulimia nervosa: Three integrated case control comparisons. *American Journal of Psychiatry*, 151, 402-407.
- Wildes, J. E., & Emery, R. E. (2001). The roles of ethnicity and culture in the development of eating disturbance and body dissatisfaction: A meta-analytic review. *Clinical Psychological Review*, 21, 521-551.

- Wilfley, D. E., Schwartz, M. B., Spurrell, E. B., & Fairburn, C. G. (1997). Assessing the specific psychopathology of binge eating disorder patients: Interview or self-report? *Behaviour Research and Therapy*, *35*, 1151-1159.
- Wilfley, D. E., Schwartz, M. B., Spurrell, E. B., & Fairburn, C. G. (2000). Using the Eating Disorder Examination to identify the specific psychopathology of binge eating disorder. *International Journal of Eating Disorders*, *27*, 259-269.
- Williamson, D. A., Gleaves, D. H., & Savin, S. M. (1992). Empirical classification of eating disorder not otherwise specified: Support for DSM-IV changes. *Journal of Psychopathology*, *14*, 201-216.
- Williamson, D. A., Womble, L. G., Smeets, M. A. M., Netemeyer, R. G., Thaw, J. M., Kutlesic, V. et al. (2002). Latent structure of eating disorder symptoms: A factor analytic and taxometric investigation. *American Journal of Psychiatry*, *159*, 412-418.
- Williamson, D. A., Zucker, N. L., Martin, C. K., & Smeets, M. (2001). Etiology and management of eating disorders. In P. B. Sutker & H. E. Adams (Eds.), *Comprehensive handbook of psychopathology (3rd ed.)* (pp. 641-670). New York, NY: Kluwer Academic/Plenum Publishers.
- Wilson, G. T., & Eldredge, K. L. (1991). Frequency of binge eating in bulimic patients: Diagnostic validity. *International Journal of Eating Disorders*, *10*, 557.

APPENDIX A
ANOREXIA NERVOSA CODING SHEET

Item (Variable label)			
Article Number			List:
Number of Citations			List:
Diagnostic Type			1 AN 2 BN 3 BED
Reviewer			1 'RR' 2 'KA' 3 'CF' 4 'MM'
Journal			List:
Authors			List:
Year			List:
Study Type			1 'Assessment/Descriptive' 2 'Treatment/Intervention' 3 'Other'
Subjects diagnostic status			1 all of AN group considered to have current AN dx 2 mixed – some current, some lifetime/past dx 3 all of AN group considered to have lifetime/past dx 4 not specified or unclear
Investigatory Department			1 psychiatry 2 psychology 3 other
SAMPLE CHARACTERISTICS:			
Number of AN subjects		n	1 'Provided for AN group' List: 2 'not mentioned at all' (for any AN group) 3 'unclear / unsure' 4 'provided for only one subtype'
AN Age mean		n	1 'Provided for AN group, List: 2 'not mentioned at all' (for any AN group) 3 'provided ONLY for AN subtype(s)'
AN Age Std Deviation		n	1 'Provided for AN group, List: 2 'not mentioned at all' (for any AN group) 3 'provided ONLY for AN subtype(s)'
AN Age Minimum		n	1 'Provided for AN group, List: 2 'not mentioned at all' (for any AN group) 3 'provided ONLY for AN subtype(s)'
AN Age Maximum		n	1 'Provided for AN group, List: 2 'not mentioned at all' (for any AN group) 3 'provided ONLY for AN subtype(s)'
Number of AN females		n	1 'Provided for AN group, List: 2 'not mentioned at all' (for any AN group) 3 'provided ONLY for AN subtype(s)'
Number of AN males		n	1 'Provided for AN group, List: 2 'not mentioned at all' (for any AN group) 3 'provided ONLY for AN subtype(s)'

RACE REPORTED?			
Were non-Caucasians definitely used?			1 'yes – other races specifically mentioned in AN group' 2 'no – only Caucasians were used in AN group' 3 'race information not provided at all' (for any AN group)
Number of AN Caucasians used			1 'Provided for AN group, List: 2 'Cauc mentioned, but number of Cauc not provided' 3 'race information provided, but Cauc not mentioned' 4 'race information not provided at all'
Number of AN African-Americans used		n	1 'Provided for AN group, List: 2 'AA mentioned, but number of AA not provided' 3 'race information provided, but AA not mentioned' 4 'race information not provided at all'
Number of AN Hispanic American used		n	1 'Provided for AN group, List: 2 'HA mentioned, but number of HA not provided' 3 'race information provided, but HA not mentioned' 4 'race information not provided at all'
Number of AN Native American used			1 'Provided for AN group, List: 2 'NA mentioned, but number of NA not provided' 3 'race information provided, but NA not mentioned' 4 'race information not provided at all'
Number of AN Asian used			1 'Provided for AN group, List: 2 'Asian mentioned, but number of Asian not provided' 3 'race information provided, but Asian not mentioned' 4 'race information not provided at all'
Number of AN 'others' for race used		n	1 'Provided for AN group, List: 2 'Other mentioned, but number of other not provided' 3 'race information provided, but other not mentioned' 4 'race information not provided at all'
SES assessed		n	1 'yes, SES provided for AN group' 2 'SES not mentioned at all' (for any AN group) 3 'SES provided, but ONLY for AN subtype'
ANBP Characteristics			
Num ANBP (binge/purge type)			1 'Provided for ANBP group, List: 2 'not mentioned at all' (for any AN group) 3 'not provided for subtype 4 'subtype not used'
ANBP Age mean			1 'Provided for ANBP group, List: 2 'not mentioned at all' (for any AN group) 3 'not provided for subtype 4 'subtype not used'
ANBP Age Std Deviation			1 'Provided for ANBP group, List: 2 'not mentioned at all' (for any AN group) 3 'not provided for subtype 4 'subtype not used'
ANBP Age min			1 'Provided for ANBP group, List: 2 'not mentioned at all' (for any AN group) 3 'not provided for subtype 4 'subtype not used'

ANBP Age max			1 'Provided for ANBP group, List: 2 'not mentioned at all' (for any AN group) 3 'not provided for subtype 4 'subtype not used'
ANBP Num of females			1 'Provided for ANBP group, List: 2 'not mentioned at all' (for any AN group) 3 'not provided for subtype 4 'subtype not used'
ANBP Num of males			1 'Provided for ANBP group, List: 2 'not mentioned at all' (for any AN group) 3 'not provided for subtype 4 'subtype not used'
RACE			
ANBP Non-Caucasians definitely used?			1 'yes –Non-Caucasians used' (in ANBP) 2 'no – only Caucasians used (in ANBP) 3 'race not mentioned at all' (for any AN group) 4 'Non-cauc not provided for subtype' (in ANBP) 5 'subtype not used'
Num of ANBP Caucasians used			1 'Provided for ANBP, List: 2 'Cauc mentioned but num Cauc not provided (in ANBP) 3 'race infor provided, but Cauc not mentioned (in ANBP) 4 'race info not provided, at all (for any AN group) 5 'subtype not used'
Num of ANBP African-Americans used			1 'Provided for ANBP, List: 2 'AA mentioned but num AA not provided (in ANBP) 3 'race infor provided, but AA not mentioned (in ANBP) 4 'race info not provided, at all (for any AN group) 5 'subtype not used'
Num of ANBP Hispanic-Americans used			1 'Provided for ANBP, List: 2 'HA mentioned but num HA not provided (in ANBP) 3 'race infor provided, but HA not mentioned (in ANBP) 4 'race info not provided, at all (for any AN group) 5 'subtype not used'
Num of ANBP Native Americans used			1 'Provided for ANBP, List: 2 'NA mentioned but num NA not provided (in ANBP) 3 'race infor provided, but NA not mentioned (in ANBP) 4 'race info not provided, at all (for any AN group) 5 'subtype not used'
Num of ANBP Asian used			1 'Provided for ANBP, List: 2 'Asian mentioned but num Asian not provided (in ANBP) 3 'race info provided, but Asian not mentioned (in ANBP) 4 'race info not provided, at all (for any AN group) 5 'subtype not used'
Num of ANBP 'others' for race used			1 'Provided for ANBP, List: 2 'Other mentioned but num other not provided (in ANBP) 3 'race infor provided, but other not mentioned (in ANBP) 4 'race info not provided, at all (for any AN group) 5 'subtype not used,
ANBP SES assessed			1 'yes – ANBP SES assessed 2 'SES not mentioned at all' (for any AN group) 3 'ANBP subtype used, but ANBP SES not provided' 4 'subtype not used'

ANR Characteristics			
Num ANR (restricting type)			1 'Provided for ANR group, List: 2 'not mentioned at all' (for any AN group) 3 'not provided for subtype 4 'subtype not used'
ANR Age mean			1 'Provided for ANR group, List: 2 'not mentioned at all' (for any AN group) 3 'not provided for subtype 4 'subtype not used'
ANR Age std Deviation			1 'Provided for ANR group, List: 2 'not mentioned at all' (for any AN group) 3 'not provided for subtype 4 'subtype not used'
ANR Age min			1 'Provided for ANR group, List: 2 'not mentioned at all' (for any AN group) 3 'not provided for subtype 4 'subtype not used'
ANR Age max			1 'Provided for ANR group, List: 2 'not mentioned at all' (for any AN group) 3 'not provided for subtype 4 'subtype not used'
ANR Num of females			1 'Provided for ANR group, List: 2 'not mentioned at all' (for any AN group) 3 'not provided for subtype 4 'subtype not used'
ANR Num of males			1 'Provided for ANR group, List: 2 'not mentioned at all' (for any AN group) 3 'not provided for subtype 4 'subtype not used'
RACE			
Were non-Caucasians ANR definitely used?			1 'yes –Non-Caucasians used' (in ANR) 2 'no – only Caucasians used (in ANR)' 3 'race not mentioned at all' (for any AN group) 4 'Non-Cauc not provided for subtype' (ANR) 5 'subtype not used'
Num of ANR Caucasians used			1 'Provided for ANR, List: 2 'Cauc mentioned but num Cauc not provided (in ANR) 3 'race infor provided, but Cauc not mentioned (in ANR)' 4 'race info not provided, at all (for any AN group) 5 'subtype not used'
Num of ANR African-Americans used			1 'Provided for ANR, List: 2 'AA mentioned but num AA not provided (in ANR) 3 'race info provided, but AA not mentioned (in ANR)' 4 'race info not provided, at all (for any AN group) 5 'subtype not used'
Num of ANR Hispanic-Americans used			1 'Provided for ANR, List: 2 'HA mentioned but num HA not provided (in ANR) 3 'race infor provided, but HA not mentioned (in ANR)' 4 'race info not provided, at all (for any AN group) 5 'subtype not used'

Num of ANR Native Americans used			1 'Provided for ANR, List: 2 'NA mentioned but num NA not provided (in ANR) 3 'race infor provided, but NA not mentioned (in ANR) 4 'race info not provided, at all (for any AN group) 5 'subtype not used
Num of ANR Asian used			1 'Provided for ANR, List: 2 'Asian mentioned but num Asian not provided (in ANR) 3 'race infor provided, but Asian not mentioned (in ANR) 4 'race info not provided, at all (for any AN group) 5 'subtype not used
Num of ANR 'others' for race used			1 'Provided for ANR, List: 2 'Other mentioned but num other not provided (in ANR) 3 'race infor provided, but other not mentioned (in ANR) 4 'race info not provided, at all (for any AN group) 5 'subtype not used
ANR SES assessed			1 'yes – ANR SES assessed 2 'SES not mentioned at all' (for any AN group) 3 'ANR subtype used, but ANR SES not provided' 4 'subtype not used

DIAGNOSTIC METHODS			
Pre-existing Diagnosis?			CIRCLE ONLY ONE 1 'Pre-existing dx <i>confirmed by researchers (additional criteria)</i> 2 'New Dx made by researchers' 3 'Pre-existing dx <i>but not confirmed (no additional criteria)</i> 4 'Mixed - some had pre-existing, some are just now dx'd' 5 'Unsure'
Diagnostic tool used (how diagnosis is defined)			1 'DSM III (1980)' 2 'DSM IV (1994) or 'DSM IV TR (2000)' 3 'ICD-10' 4 'No tool mentioned' 5 'Other' List: (also, if two or more tools are mentioned)
Diagnostic methods reported			1 'yes – method used to make diagnosis was reported' 2 'no'
All participants dx'd same way			1 'yes 2 'no 3 'unsure
Dx – structured interview with individual			1 'yes' List: 2 'diagnostic methods reported, but struct interview not used' 3 'diagnostic methods not reported'
Dx – structured interview with other informant			1 'yes' List: 2 'diagnostic methods reported, but struct int w/other not used' 3 'diagnostic methods not reported' Other - identify -

Dx – semi-structured interview with individual (e.g., “clinical interview”)			1 ‘yes’ List: 2 ‘diagnostic methods reported, but semi-struct int not used’ 3 ‘diagnostic methods not reported’
Dx – semi-structured interview with other (e.g., “clinical interview”)			1 ‘yes’ List: 2 ‘diag methods reported, but semi-struct int w/other not used’ 3 ‘diagnostic methods not reported’ Other - identify -
Dx – type of questionnaire used (if used for diagnosis)			1 ‘EDE-Q’ 2 ‘EAT’ 3 ‘EDQ’ 4 ‘dx reported, but questionnaires not used for diagnosis’ 5 ‘diagnostic methods not reported’ 6 ‘other’ List:
Referral or recruitment source specified?			1 ‘in-patient’ 2 ‘out-patient’ 3 ‘mixed’ 4 ‘community sample’ 5 ‘not stated / unsure’

CRITERIA DEFINITIONS			
AMENORRHEA			
Amenorrhea specifically required (for three consecutive months)			1 ‘yes – amenorrhea specifically stated’ 2 ‘no – amenorrhea specifically not required for inclusion’ 3 ‘dx assessed through interview/questionnaire, and amenorrhea not otherwise mentioned’ 4 ‘amenorrhea not specifically mentioned at all’
BODY WEIGHT			
Body weight criterion specifically required (<85% of expected weight)			1 ‘yes – body weight criterion specifically stated’ 2 ‘no – body weight criterion specifically not required for inclusion’ 3 ‘dx assessed through interview/questionnaire, and body weight criterion not otherwise mentioned’ 4 ‘body weight criterion not mentioned at all’
Body weight assessed through			1 ‘BMI’ 2 ‘American Insurance Industry’s <i>Build and Blood Pressure Study</i> 3 ‘Metropolitan Life Company weight charts’ 4 ‘Lbs’ 5 ‘Other’ List: 6 ‘Not specifically stated’
Minimum BMI/Weight cut-off for participation			1 ‘yes’ List: 2 ‘none reported’

Maximum BMI/Weight cut off for participation			1 'yes' List: 2 'none reported'
Actual time period of eating disturbance reported			1 'yes' 2 'no'
<i>SUBTYPES:</i>			
AN Subtypes specifically assessed			1 'yes – subtypes were assessed, and subtypes groups were included/compared' 2 'yes – subtypes were assessed, but subtype groups were not included/compared' 3 'no – subtypes were not assessed at all'
<i>ANBP BINGE:</i>			
Binge definition provided for diagnosis			1 'yes – specific binge definition was provided for dx/ inclusion (objective/subject/social-circum.) in ANBP' 2 'diagnosis/binging was assessed through interview or questionnaire and <u>binge definition otherwise not mentioned</u> ' 3 'binge assessment was just not reported (dx was not assessed through interview/questionnaire) 4 'ANBP subtype not used'
Binges required to have met caloric or objective standard in order to count as a binge (for dx)			1 'yes – List (cut-off or minimum required): 2 'no – binge criteria was described, but objective standard was specifically not required' (ANBP used) 3 'dx was assessed through interview/questionnaire, <u>binge def otherwise not mentioned</u> ' (ANBP used) 4 'binge assessment just not reported' (ANBP used) 5 'ANBP subtype not used'
Binges defined subjectively by part. (for dx)			1 'yes – binges subjectively reported and no binge criteria was used (ANBP used) 2 'no –binge criteria was described, but subjective report was not used' (ANBP used) 3 'dx was assessed through interview/questionnaire, <u>binge def otherwise not mentioned</u> ' (ANBP used) 4 'binge assessment just not reported' (ANBP used) 5 'ANBP subtype not used'
Binges defined 'social-circumstantially' (for dx)			1 'yes –binges reported social-circumstantially' (ANBP used) 2 'no –binge criteria was described, but social-circumstance report was not used' (ANBP used) 3 ' dx was assessed through interview/questionnaire, <u>binge def otherwise not mentioned</u> ' (ANBP used) 4 'binge assessment just not reported' (ANBP used) 5 'ANBP subtype not used'
Minimum binge frequency required for ANBP			1 'yes – required to meet minimum frequency. (ANBP used) List: 2 'no –binge criteria described, but no minimum frequency was required' (ANBP used) 3 'dx was assessed through interview/questionnaire, binge frequency otherwise not mentioned' (ANBP used) 4 'binge frequency just not reported' (ANBP used) 5 'ANBP subtype not used'

Lack of control over eating specifically required for binge in ANBP			1 'yes – lack of control required for binge' (ANBP used) 2 'no – lack of control not required' (ANBP used) 3 'dx was assessed through interview/questionnaire, <u>binge def otherwise not mentioned</u> ' (ANBP used) 4 'binge assessment just not reported' (ANBP used) 5 'ANBP subtype not used'
COMPENSATORY BX:			
Minimum compensatory bx freq (cut off) required for inclusion (e.g., 2x/wk for 3 months) (for diagnosis)			1 'yes – minimum compensatory behavior frequency required' (List if different from 2x/wk for 3 month) : 2 'no – minimum comp bx frequency specifically not required' 3 'dx was assessed through interview/questionnaire, and <u>compensatory behavior otherwise not mentioned</u> ' 4 'compensatory bx assessment just not reported'
Compensatory behaviors specifically assessed			1 'yes' 2 'dx assessed through interview/questionnaire and <u>compensatory behavior otherwise not mentioned</u> ' 3 'compensatory behavior not mentioned or reported at all'
Type of Compensatory behavior reportedly assessed			1 'only purging/vomit' (if sample is ANBP only) 2 'only laxative / diuretic / enema abuse' (if sample is ANBP only) 3 'only exercise' (if sample is ANR only) 4 'only fasting/restricting' (if sample is ANR only) 5 'multiple compensatory behaviors – List: ' 6 'dx assessed through interview/questionnaire and <u>compensatory behavior otherwise not mentioned</u> ' 7 'compensatory behavior just not assessed or reported'

DUAL DIAGNOSIS			
AN as only Axis I or Axis II dx allowed?			1 'Yes - AN only was specifically stated' 2 'No - Dual or Mixed diagnoses specifically allowed' 3 'Unsure - not specifically stated'
OCD specifically permitted			1 'Yes - ocd also present' 2 'No'
Substance abuse specifically permitted			1 'Yes – substance abuse also present' 2 'No'
Personality d/os specifically permitted			1 'Yes – personality disorders also present' 2 'No'
Borderline Pers. DO specifically permitted			1 'Yes – borderline disorder also present' 2 'No'
Bipolar Pers DO specifically permitted			1 'Yes – bipolar personality disorder also present' 2 'No'

Mood/depressive disorders specifically permitted			1 'Yes - mood/depr disorders also present' 2 'No'
Anxiety disorders specifically permitted			1 'Yes - anxiety disorders also present' 2 'No'
Other diagnoses specifically permitted			1 'Yes - some other dx (besides above) permitted' (List here) 2 'No'
Hx of other ED specifically permitted			1 'yes' 2 'no'
Medical problems assessed as part of the subject selection process			1 'yes' 2 'no' (ignore if medical problems followed only as a dependent or outcome variable)

EXCLUSIONARY CRITERIA			
Exclusionary criteria Mentioned			1 'Yes - exclusionary criteria (i.e., any other diagnosis) was mentioned or described' 2 'No'
OCD specifically excluded?			1 'Yes - OCD excluded' 2 'No - OCD not excluded'
Substance abuse specifically excluded?			1 'Yes - substance abuse excluded' 2 'No - substance abuse not excluded'
Personality disorders specifically excluded?			1 'Yes - personality disorders excluded' 2 'No - personality disorders not excluded'
Bipolar disorders specifically excluded?			1 'Yes - Bipolar disorders excluded' 2 'No - bipolar disorders not excluded'
Borderline disorder specifically excluded?			1 'Yes - Borderline disorder excluded' 2 'No - Borderline disorder not excluded'
Avoidant PD specifically excluded?			1 'Yes - avoidant p. disorder specifically excluded' 2 'No - avoidant personality disorder not excluded'
Mood/Depressive disorders specifically excluded?			1 'Yes - mood/depressive disorders excluded' 2 'No - mood/depressive disorders not excluded'

Anxiety Disorders specifically excluded?			1 'Yes - anxiety disorders excluded' 2 'No - anxiety disorders not excluded'
Psychotic disorders specifically excluded?			1 'Yes - psychotic disorders excluded' 2 'No - psychotic disorders not excluded'
Obese individuals specifically excluded?			1 'Yes - obese excluded' 2 'No - obese not excluded'
Hx of Other eating disorder specifically excluded			1 'Yes' List: 2 'No - history of other eating disorders not specifically excluded'
Zero evidence that Dual/Extra Diagnoses were considered?			1 'Yes - Dual diagnoses were simply not mentioned or referred to in any way' 2 'No - dual diagnoses were mentioned somehow (whether included or excluded)'
Other criteria specifically excluded			1 'Yes - some other criteria (besides above - e.g., no hx of prior med use/tx) excluded' (List:) 2 'No'
COMORBID HEALTH CONDITIONS			
List other medical problems tracked in patients with AN			List:

APPENDIX B
BULIMIA NERVOSA CODING SHEET

Item (Variable label)	Variable Name	Type	Values
Reviewer	Reviewer	c	1 'RR' 2 'KA' 3 'CF' 4 'MM'
Date of Review	Date	date	
Journal	Journal	Jnl	
Authors	Author	s	
Title	Title	s	
Year	Year	s	
Vol/pp	Vol&pp	s	
Study Type	Studytyp	s	1 'Assessment/Descriptive' 2 'Treatment/Intervention' 3 'Other'
Subjects diagnostic status			1 all of BN group considered to have current BN dx 2 mixed – some current, some lifetime/past dx 3 all of BN group considered to have lifetime/past dx 4 not specified or unclear

SAMPLE CHARACTERISTICS:			
Where BN = BNP + BNNP			
Number of BN subjects			1 'Provided for BN group, List: 2 'not mentioned at all' (for any BN group) 3 'unclear / unsure'
BN Age mean			1 'Provided for BN group, List: 2 'not mentioned at all' (for any BN group)
BN Age Std Deviation			1 'Provided for BN group, List: 2 'not mentioned at all' (for any BN group)
BN Age Minimum			1 'Provided for BN group, List: 2 'not mentioned at all' (for any BN group)
BN Age Maximum			1 'Provided for BN group, List: 2 'not mentioned at all' (for any BN group)
Number of BN females			1 'Provided for BN group, List: 2 'not mentioned at all' (for any BN group)
Number of BN males			1 'Provided for BN group, List: 2 'not mentioned at all' (for any BN group)
RACE REPORTED?			
Were non-Caucasians definitely used?			1 'yes – other races specifically mentioned in BN group' 2 'no – only Caucasians were used in BN group' 3 'race information provided, but ONLY by BN subtype' 4 'race information not provided at all' (for any BN group)
Number of BN African-Americans used			1 'Provided for BN group, List: 2 race info 'provided, but ONLY divided by BN subtype' 3 'AA mentioned for BN but num of AA no provided' 4 'race info provided , but num of AA not mentioned' 5 'race not mentioned at all' (for any BN group)
Number of BN Hispanic American used			1 'Provided for BN group, List: 2 'race info provided, but ONLY divided by BN subtype' 3 'HA mentioned for BN but num of HA not provided' 4 'race info provided, but num of HA not mentioned' 5 'race not mentioned at all' (for any BN group)

Number of BN Asian used			1 'Provided for BN group, List: 2 'race info provided, but ONLY divided by BN subtype' 3 'Asian mentioned for BN but num of Asian not provided' 4 'race info provided, but num of Asian not mentioned' 5 'race not mentioned at all' (for any BN group)
Number of BN 'others' for race used			1 'Provided for BN group, List: 2 'race info provided, but ONLY divided by BN subtype' 3 'other' mentioned for BN but num of other not provided' 4 'race info provided, but num of "other" not mentioned' 5 'race not mentioned at all' (for any BN group)
SES assessed			1 'yes, SES provided for BN group'' 2 'provided, but ONLY divided by BN subtype' 3 'SES not mentioned at all' (for any BN group)

SUBTYPE CHARACTERISTIC (use ONLY if BN described by either/both subtypes BN-P AND BN-NP)			
Num BN-P (bulimia nervosa – purging type)			1 'BNP group included, provided for BNP, List: 2 'subject number not mentioned at all' (for any BN group) 3 'BNP subtype used, but sub number <u>not</u> provided for BNP' 4 'BNP subtype not used but sub num described in BNNP only' 5 'BNP subtype not used but sub num described in BN only'
BN-P Age mean			1 'BNP group included, provided for BNP, List: 2 'mean not mentioned at all' (for any BN group) 3 'BNP subtype used, but mean <u>not</u> provided for BNP' 4 'BN subtype not used but mean described in BNNP only' 5 'BN subtype not used but mean described in BN only'
BN-P Age Std Deviation			1 'BNP group included, provided for BNP, List: 2 'std deviation not mentioned at all' (for any BN group) 3 'BNP subtype used, but std deviation <u>not</u> provided for BNP' 4 'subtype not used but std dev described in BNNP only' 5 'subtype not used but std dev described in BN only'
BN-P Age min			1 'BNP group included, provided for BNP, List: 2 'age min not mentioned at all' (for any BN group) 3 'BNP subtype used, but age min <u>not</u> provided for BNP' 4 'subtype not used but age min described in BNNP only' 5 'subtype not used but age min described in BN only'
BN-P Age max			1 'BNP group included, provided for BNP, List: 2 'age max not mentioned at all' (for any BN group) 3 'BNP subtypes used, but max age <u>not</u> provided for BNP' 4 'subtype not used but age max described in BNNP only' 5 'subtype not used, but age max described in BN only'
BN-P Num of females			1 'BNP group included, provided for BNP, List: 2 'number of females not mentioned at all' (for any BN group) 3 'BNP subtype used, but num of fems not provided for BNP' 4 'subtype not used but num of fems described in BNNP only' 5 'subtype not used but num of fems described in BN only'

BN-P Num of males			1 'BNP group included, provided for BNP, List: 2 'number of males not mentioned at all' (for any BN group) 3 'BNP subtypes used, but number of BNP males not 4 'subtype not used but num of males described in BNNP only' 5 'subtype not used but num of males described in BN only'
BN-P Non-Caucasians definitely used?			1 'yes – BNP Non-Caucasians used' 2 'no – only Caucasians used in BNP group' 3 'race not mentioned at all' (for any BN group) 4 'BNP subtype used, but non-Caucs for BNP not provided' 5 'subtype not used but race info described in BNNP only' 6 'subtype not used but race info described in BN only'
Num of BN-P African-Americans used			1 'BNP group included, provided for BNP, List: 2 'BNP subtype used, but num of BNP A-Ams not provided' 3 'subtype not used but num of A-As described in BNNP only' 4 'subtype not used but num of A-Ams described in BN only' 5 'race info provided, but num of AA BNP not mentioned' 6 'race info not mentioned at all' (for any BN group)
Num of BN-P Hispanic-Americans used			1 'BNP group included, provided for BNP, List: 2 'BNP subtype used, but num of BNP His-Ams not provided' 3 'subtype not used but num of H-As described in BNNP only' 4 'subtype not used but num of H-A described in BN only' 5 'race info provided, but num of HA BNP not mentioned' 6 'race not mentioned at all' (for any BN group)
Num of BN-P Asian used			1 'BNP group included, provided for BNP, List: 2 'BNP subtype used, but num of BNP Asian not provided' 3 'subtype not used but num of Asian described in BNNP only' 4 'subtype not used but num of Asian described in BN only' 5 'race info provided, but num of Asian BNP not mentioned' 6 'race not mentioned at all' (for any BN group)
Num of BN-P 'others' for race used			1 'BNP group included, provided for BNP, List: 2 'BNP subtype used, but num of BNP His-Ams not provided' 3 'subtype not used but num of other described in BNNP only' 4 'subtype not used but num of other described in BN only' 5 'race info provided, but num of other BNP not mentioned' 6 'race not mentioned at all' (for any BN group)
BN-P SES assessed			1 'yes – BNP SES assessed' 2 'SES not mentioned at all' (for any BN group) 3 'BNP subtype used, but BNP SES not provided' 4 'subtype not used but SES described in BNNP only' 5 'subtype not used but SES described in BN only'
Num BN-NP (bulimia nervosa, non-purge type)			1 'BNNP group included, provided for BNNP, List: 2 'subject number not mentioned at all' (for any BN group) 3 'BNNP subtype used, but sub num <u>not</u> provided for BNNP' 4 'subtype not used but sub num described in BNP only' 5 'subtype not used but sub num described in BN only'

BN-NP Age mean			1 'BNNP group included, provided for BNNP, List: 2 'mean not mentioned at all' (for any BN group) 3 'BNNP subtype used, but mean <u>not</u> provided for BNNP 4 'subtype not used but mean described in BNP only' 5 'subtype not used but mean described in BN only'
BN-NP Age std Deviation			1 'BNNP group included, provided for BNNP, List: 2 'std deviation not mentioned at all' (for any BN group) 3 'BNNP subtype used, but std dev <u>not</u> provided for BNNP 4 'subtype not used but std dev described in BNP only' 5 'subtype not used but std dev described in BN only'
BN-NP Age min			1 'BNNP group included, provided for BNNP, List: 2 'age min not mentioned at all' (for any BN group) 3 'BNNP subtype used, but age min <u>not</u> provided for BNNP 4 'subtype not used but age min described in BNP only' 5 'subtype not used but age min described in BN only'
BN-NP Age max			1 'BNNP group included, provided for BNNP, List: 2 'age max not mentioned at all' (for any BN group) 3 'BNNP subtypes used, but max age <u>not</u> provided for BNNP 4 'subtype not used but age max described in BNP only' 5 'subtype not used but age max described in BN only'
BN-NP Num of females			1 'BNNP group included, provided for BNNP, List: 2 'number of females not mentioned at all' (for any BN group) 3 'BNNP subtype used, but num of BNNP fems not provided' 4 'subtype not used but num of fems described in BNP only' 5 'subtype not used but num of fems described in BN only'
BN-NP Num of males			1 'BNNP group included, provided for BNNP, List: 2 'number of males not mentioned at all' (for any BN group) 3 'BNNP subtypes used, but num of BNNP males not provided' 4 'subtype not used but num of males described in BNP only' 5 'subtype not used but num of males described in BN only'
Were non-Caucasians BN-NP definitely used?			1 'yes – BNP Non-Caucasians used' 2 'no – only Caucasians used in BNNP group' 3 'race not mentioned at all' (for any BN group) 4 'BNP subtype used, but non-Cauc for BNNP not provided' 5 'subtype not used but race info described in BNNP only' 6 'subtype not used but race info described in BN only'
Num of BN-NP African-Americans used			1 'BNNP group included, provided for BNNP, List: 2 'BNNP subtype used, but num of BNNP A-Ams not provided' 3 'subtype not used but num of A-As described in BNP only' 4 'subtype not used but num of A-Ams described in BN only' 5 'race info provided, but num of AA BNNP not mentioned' 6 'race info not mentioned at all' (for any BN group)
Num of BN-NP Hispanic-Americans used			1 'BNNP group included, provided for BNNP, List: 2 'BNNP subtype used, but num of BNNP H-Ams not provided' 3 'subtype not used but num of Has described in BNP only' 4 'subtype not used but num of H-Ams described in BN only' 5 'race info provided, but num of HA BNNP not mentioned' 6 'race info not mentioned at all' (for any BN group)

Num of BN-NP Asian used			1 'BNNP group included, provided for BNNP, List: 2 'BNNP subtype used, but num of BNNP Asian not provided' 3 'subtype not used but num of Asian described in BNP only' 4 'subtype not used but num of Asian described in BN only' 5 'race info provided, but num of Asian BNNP not mentioned' 6 'race info not mentioned at all' (for any BN group)
Num of BN-NP 'others' for race used			1 'BNNP group included, provided for BNNP, List: 2 'BNNP subtype used, but num of BNNP other not provided' 3 'subtype not used but num of other described in BNP only' 4 'subtype not used but num of other described in BN only' 5 'race info provided, but num of other BNNP not mentioned' 6 'race info not mentioned at all' (for any BN group)
BN-NP SES assessed			1 'yes – BNNP SES assessed 2 'SES not mentioned at all' (for any BN group) 3 'BNNP subtype used, but BNNP SES not provided' 4 'subtype not used but SES described in BNP only' 5 'subtype not used but SES described in BN only'

DIAGNOSTIC METHODS			
Pre-existing Diagnosis?			CIRCLE ONLY ONE 1 'Pre-existing dx <i>confirmed by researchers (additional criteria)</i> 2 'New Dx made by researchers' 3 'Pre-existing dx <i>but not confirmed (no additional criteria)</i> 4 'Mixed - some had pre-existing, some are just now dx'd' 5 'Unsure'
Diagnostic tool used (how diagnosis is defined)			1 'DSM III (1980)' 2 'DSM IV (1994) or 'DSM IV TR (2000)' 3 'ICD-10' 4 'No tool mentioned' 5 'Other' List: (also, if two or more tools are mentioned)
Diagnostic methods reported			1 'yes – method used to make diagnosis was reported' 2 'no'
Dx – structured interview with individual			1 'yes' List: 2 'diagnostic methods reported, but struct interview not used' 3 'diagnostic methods not reported'
Dx – structured interview with other informant			1 'yes' List: 2 'diagnostic methods reported, but struct int w/other not used' 3 'diagnostic methods not reported' Other - identify -
Dx – semi-structured interview with individual (e.g., "clinical interview")			1 'yes' List: 2 'diagnostic methods reported, but semi-struct int not used' 3 'diagnostic methods not reported'

Dx – semi-structured interview with other (e.g., “clinical interview”)			1 ‘yes’ List: 2 ‘diag methods reported, but semi-struct int w/other not used’ 3 ‘diagnostic methods not reported’ Other - identify -
Dx – food diaries (dx/inclusion)			1 ‘yes – food diaries used to determine diagnosis/ inclusion’ 2 ‘diagnostic methods reported, but food diaries not used’ 3 ‘diagnostic methods not reported’
Dx – laboratory feeding methods (used for dx/inclusion)			1 ‘yes – lab feeding was used to determine diagnosis/inclusion’ 2 ‘diagnostic methods reported, but lab feeding not used’ 3 ‘diagnostic methods not reported’
Dx – type of questionnaire used (if used for diagnosis)			1 ‘EDE’ 2 ‘EAT’ 3 ‘EDQ’ 4 ‘dx reported, but questionnaires not used for diagnosis’ 5 ‘diagnostic methods not reported’ 6 ‘ other’ List:
Referral or recruitment source specified?			1 ‘in-patient’ 2 ‘out-patient’ 3 ‘mixed’ 4 ‘community’ 5 ‘unsure’

CRITERIA DEFINITIONS			
<i>BINGE:</i>			
Binge definition provided for diagnosis			1 ‘yes – specific binge definition was provided for dx/ inclusion (objective/subject/social-circum.)’ 2 ‘uncertain – diagnosis/binging was assessed through structured/semi-struct interview or questionnaire and binge definition otherwise not mentioned’ 3 ‘binge assessment was just not reported (dx was not assessed through interview/questionnaire)’
Binges required to have met caloric or objective standard in order to count as a binge (for dx)			1 ‘yes’ List (cut-off or minimum required): 2 ‘no – binge criteria was described, but objective standard was specifically not required’ 3 ‘uncertain – dx was assessed through structured/semi-structured interview/questionnaire, binge def otherwise not mentioned’ 4 ‘binge assessment just not reported’
Binges defined subjectively by part. (for dx)			1 ‘yes – binges were self-reported and no binge criteria was used’ 2 ‘no – binge criteria was described, subjective definition not permitted’ 3 uncertain – dx was assessed through nterview/questionnaire and binge def otherwise not mentioned’ 4 ‘binge assessment just not reported’

Binges defined 'social-circumstantially' (for dx)			<ol style="list-style-type: none"> 1 'yes – presence of binges were assessed by circumstances (binges said to be larger amt of food than would normally be consumed)' 2 'no – binge criteria was described, but not assessed by circumstances' 3 'uncertain – dx as assessed through interview/questionnaire' 4 'binge assessment just not reported'
Number of days on which binge eating occurs counted (for dx or inclusion crit.)			<ol style="list-style-type: none"> 1 'yes – binge frequency described by number of DAYS' 2 'no – binge frequency not described by number of days' 3 'uncertain – dx was assessed through interview/questionnaire, binge frequency otherwise not mentioned' 4 'binge assessment just not reported' <p>(days with binges counted - NOT # of binges/ day)</p>
Number of episodes of binge eating counted (dx)			<ol style="list-style-type: none"> 1 'yes – binge frequency described by number of EPISODES' 2 'no – binge frequency not described by episode' 3 'uncertain – dx assessed through interview/questionnaire and binge frequency otherwise not mentioned' 4 'binge assessment just not reported' <p>(binges counted by episode - could be # of binges/day)</p>
Minimum binge freq (cut off) required for inclusion (e.g., 2x/wk for 3 months) (for diagnosis)			<ol style="list-style-type: none"> 1 'yes – minimum binge frequency required' (List if different from 2x/wk for 3 month) : 2 'no – minimum binge frequency specifically not required' 3 'uncertain – dx was assessed through interview/questionnaire and binge frequency otherwise not mentioned' 4 'binge assessment just not reported'
Lack of control over eating specifically required for binge			<ol style="list-style-type: none"> 1 'yes – lack of control specifically required' 2 'lack of control specifically not required' 3 'uncertain – dx was assessed through interview/questionnaire, lack of control otherwise not mentioned' 4 'lack of control assessment just not reported'
COMPENSATORY BX:			
Minimum compensatory bx freq (cut off) required for inclusion (e.g., 2x/wk for 3 months) (for diagnosis)			<ol style="list-style-type: none"> 1 'yes – minimum compensatory behavior frequency required' (List if different from 2x/wk for 3 month) : 2 'no – minimum comp bx frequency specifically not required' 3 'uncertain – dx was assessed through interview/questionnaire, and compensatory behavior otherwise not mentioned' 4 'compensatory bx assessment just not reported'
Type of Compensatory behavior reportedly assessed			<ol style="list-style-type: none"> 1 'only purging/vomit' (if BNP only) 2 'only laxative / diuretic / enema abuse' (if BNP only) 3 'only exercise' (if BNNP only) 4 'only fasting/restricting' (if BNNP only) 5 'multiple compensatory behaviors – List: ' 6 'uncertain – dx assessed through interview/questionnaire and compensatory beh otherwise not mentioned'' 7 'compensatory behavior just not assessed or reported'

BODY WEIGHT:			
BMI/Weight cut-off for participation			1 'yes' List: 2 'none reported'
Body weight assessed through			1 'BMI data reported descriptively' 2 'American Insurance Industry's <i>Build & Blood Pressure Study</i> 3 'Metropolitan Life Company weight charts' 4 'Lbs' 5 'Other' List: 6 'Not reported or not assessed descriptively'
BODY IMAGE DISTURBANCE/ SELF EVALUATION DISTURBANCE			
Self-evaluation disturbance assessed through			1 'uncertain, dx assessment through questionnaire and self-eval otherwise not mentioned' 2 'uncertain, dx assessment through (structure/semistruct) interview and self-eval otherwise not mentioned' 3 'self-evaluation disturbance not assessed or reported' 4 'other' List:
Actual time period of eating disturbance reported			1 'yes' 2 'no'
SUBTYPES:			
BN Subtypes specifically assessed			1 'yes – subtypes were assessed, and subtypes groups were included/compared' 2 'yes – subtypes were assessed, but subtype groups were not included/compared' 3 'no – subtypes were not assessed at all'
DUAL DIAGNOSIS			
BN as only Axis I or Axis II dx allowed?			1 'Yes - BN only was specifically stated' 2 'No - Dual or Mixed diagnoses specifically allowed' 3 'Unsure - not specifically stated'
OCD specifically permitted			1 'Yes - ocd also present' 2 'No'
Substance abuse specifically permitted			1 'Yes – substance abuse also present' 2 'No'
Personality d/os specifically permitted			1 'Yes – personality disorders also present' 2 'No'
Borderline Pers. DO specifically permitted			1 'Yes – borderline disorder also present' 2 'No'

Bipolar Pers DO specifically permitted			1 'Yes – bipolar personality disorder also present' 2 'No'
Mood/depressive disorders specifically permitted			1 'Yes - mood/depr disorders also present' 2 'No'
Anxiety disorders specifically permitted			1 'Yes - anxiety disorders also present' 2 'No'
Other diagnoses specifically permitted			1 'Yes - some other dx (besides above) permitted' (List here) 2 'No'
Hx of other ED specifically permitted			1 'yes' 2 'no'
Medical problems assessed as part of the subject selection process			1 'yes' 2 'no' (ignore if medical problems followed only as a dependent or outcome variable)

EXCLUSIONARY CRITERIA			
Exclusionary criteria Mentioned			1 'Yes - exclusionary criteria (i.e., any other diagnosis) was mentioned or described' 2 'No'
OCD specifically excluded?			1 'Yes - OCD excluded' 2 'No - OCD not excluded'
Substance abuse specifically excluded?			1 'Yes - substance abuse excluded' 2 'No - substance abuse not excluded'
Personality disorders specifically excluded?			1 'Yes – personality disorders excluded' 2 'No – personality disorders not excluded'
Bipolar disorders specifically excluded?			1 'Yes - Bipolar disorders excluded' 2 'No - bipolar disorders not excluded'
Borderline disorder specifically excluded?			1 'Yes – Borderline disorder excluded' 2 'No – Borderline disorder not excluded'
Avoidant PD specifically excluded?			1 'Yes – avoidant p. disorder specifically excluded' 2 'No – avoidant personality disorder not excluded'

Mood/Depressive disorders specifically excluded?			1 'Yes - mood/depressive disorders excluded' 2 'No - mood/depressive disorders not excluded'
Anxiety Disorders specifically excluded?			1 'Yes - anxiety disorders excluded' 2 'No - anxiety disorders not excluded'
Psychotic disorders specifically excluded?			1 'Yes - psychotic disorders excluded' 2 'No - psychotic disorders not excluded'
Obese individuals specifically excluded?			1 'Yes - obese excluded' 2 'No - obese not excluded'
Hx of Other eating disorder specifically excluded			1 'Yes' List: 2 'No – history of other eating disorders not specifically excluded'
Zero evidence that Dual/Extra Diagnoses were considered?			1 'Yes - Dual diagnoses were simply not mentioned or referred to in any way' 2 'No - dual diagnoses were mentioned somehow (whether included or excluded)'

Other criteria specifically excluded			1 'Yes - some other criteria (besides above – e.g., no hx of prior med use/tx) excluded' (List:) 2 'No'
--	--	--	--

COMORBID HEALTH CONDITIONS			
List other medical problems tracked in patients with BN			List:

APPENDIX C
BINGE-EATING DISORDER CODING SHEET

Item (Variable label)	Variable Name	Type	Values
Reviewer	Reviewer	c	1 'RR' 2 'KA' 3 'CF' 4 'MM'
Date of Review	Date	date	
Journal	Journal	Jnl	
Authors	Author	s	
Title	Title	s	
Year	Year	s	
Vol/pp	Vol&pp	s	
Study Type	Studytype	s	1 'Assessment/Descriptive' 2 'Treatment/Intervention' 3 'Other'
Subjects diagnostic status			1 all of BED group considered to have current BED dx 2 mixed – some current, some lifetime/past dx 3 all of BED group considered to have lifetime/past dx 4 not specified or unclear
Investigatory Department			1 'psychiatry' 2 'psychology' 3 'other/unsure'

SAMPLE CHARACTERISTICS			
Number of BED subjects		n	1 'Provided for BN group, List: 2 'not mentioned at all' 3 'unclear / unsure'
BED Age mean		n	1 'Provided for BED group, List: 2 'not mentioned at all'
BED Age Std Deviation		n	1 'Provided for BED group, List: 2 'not mentioned at all'
BED Age Minimum		n	1 'Provided for BED group, List: 2 'not mentioned at all'
BED Age Maximum		n	1 'Provided for BED group, List: 2 'not mentioned at all'
Number of BED females		n	1 'Provided for BED group, List: 2 'not mentioned at all'
Number of BED males		n	1 'Provided for BED group, List: 2 'not mentioned at all'
RACE REPORTED?			
Were non-Caucasians definitely used?			1 'yes – other races specifically mentioned in BED group' 2 'no – only Caucasians were used in BED group' 3 'race information not provided at all'
Number of BED Caucasians used			1 'Provided for BED group, List: 2 'Cauc mentioned, but number of Cauc not provided' 3 'race information provided, but Cauc not mentioned' 4 'race information not provided at all'

Number of BED African-Americans used		n	1 'Provided for BED group, List: 2 'AA mentioned, but number of AA not provided' 3 'race information provided, but AA not mentioned' 4 'race information not provided at all'
Number of BED Hispanic American used		n	1 'Provided for BED group, List: 2 'HA mentioned, but num of HA not provided' 3 'race information provided, but HA not mentioned' 4 'race information not mentioned at all'
Number of BED Native American Used		N	1 'Provided for BED group, List: 2 'NA mentioned, but num of NA not provided' 3 'race information provided, but NA not mentioned' 4 'race information not mentioned at all'
Number of BED 'others' for race used		n	1 'Provided for BED group, List: 2 'other race mentioned, but num of other race not provided' 3 'race information provided, but other race not mentioned' 4 'race information not mentioned at all'
SES assessed		N	1 'yes, SES provided for BED group' 2 'SES not mentioned at all'

DIAGNOSTIC METHODS			
Pre-existing Diagnosis?		c	CIRCLE ONLY ONE 1 'Pre-existing dx <i>confirmed by researchers (additional criteria)</i> 2 'New Dx made by researchers' 3 'Pre-existing dx <i>but not confirmed (no additional criteria)</i> 4 'Mixed - some had pre-existing, some are just now dx'd' 5 'Unsure'
Diagnostic tool used (how diagnosis is defined)			1 'DSM IV (1994) or 'DSM IV TR (2000)' 2 'No tool mentioned' 3 'Other' List: (also, if two or more tools are mentioned)
Diagnostic methods reported			1 'yes – method used to make diagnosis was reported' 2 'no'
Dx – structured interview with individual			1 'yes' List: 2 'diagnostic methods reported, but struct interview not used' 3 'diagnostic methods not reported'
Dx – structured interview with other informant			1 'yes' List: 2 'diagnostic methods reported, but struct int w/other not used' 3 'diagnostic methods not reported' Other - identify
Dx – semi-structured interview with individual (e.g., "clinical interview")			1 'yes' List: 2 'diagnostic methods reported, but semi-struct int not used' 3 'diagnostic methods not reported'

Dx – semi-structured interview with other (e.g., “clinical interview”)			1 ‘yes’ List: 2 ‘diagnostic methods reported, but semi-struct int w/other not used’ 3 ‘diagnostic methods not reported’ Other - identify
Dx – food diaries (dx/ inclusion)			1 ‘yes – food diaries used to determine diagnosis/ inclusion’ 2 ‘diagnostic methods reported, but food diaries not used’ 3 ‘diagnostic methods not reported’
Dx – type of questionnaire used (if used for diagnosis)			1 ‘EDE’ 2 ‘EAT’ 3 ‘EDQ’ 4 ‘dx reported, but questionnaires not used for diagnosis’ 5 ‘diagnostic methods not reported’ 6 ‘other’ List:
Referral or recruitment source specified?			1 ‘in-patient’ 2 ‘out-patient’ 3 ‘mixed’ 4 ‘not stated / unsure’

CRITERIA DEFINITIONS			
<i>BINGE:</i>			
Binge definition provided for diagnosis			1 ‘yes – specific binge definition was provided for dx/ inclusion (objective/subject/social-circum.)’ 2 ‘uncertain – diagnosis/binging was assessed through structured/semi-struct interview or questionnaire and binge definition otherwise not mentioned’ 3 ‘binge assessment was just not reported (dx was not assessed through interview/questionnaire)’
Binges required to have met caloric or objective standard in order to count as a binge (for dx)			1 ‘yes’ List (cut-off or minimum required): 2 ‘no – binge criteria was described, but objective standard was specifically not required’ 3 ‘uncertain – dx was assessed through structured/semi-structured interview/questionnaire, binge def otherwise not mentioned’ 4 ‘binge assessment just not reported’
Binges defined subjectively by part. (for dx)			1 ‘yes – binges were self-reported and no binge criteria was used’ 2 ‘no – binge criteria was described, subjective definition not permitted’ 3 uncertain – dx was assessed through interview/questionnaire and binge def otherwise not mentioned’ 4 ‘binge assessment just not reported’

Binges defined 'social-circumstantially' (for dx)			<ul style="list-style-type: none"> 1 'yes – presence of binges were assessed by circumstances (binges said to be larger amt of food than would normally be consumed)' 2 'no – binge criteria was described, but not assessed by circumstances' 3 'uncertain – dx as assessed through interview/questionnaire' 4 'binge assessment just not reported'
Number of days on which binge eating occurs counted (for dx or inclusion crit.)			<ul style="list-style-type: none"> 1 'yes – binge frequency described by number of DAYS' 2 'no – binge frequency not described by number of days' 3 'uncertain – dx was assessed through interview/ questionnaire, binge frequency otherwise not mentioned' 4 'binge assessment just not reported' <p>(days with binges counted - NOT # of binges/ day)</p>
Number of episodes of binge eating counted (dx)			<ul style="list-style-type: none"> 1 'yes – binge frequency described by number of EPISODES' 2 'no – binge frequency not described by episode' 3 'uncertain – dx assessed through interview/questionnaire and binge frequency otherwise not mentioned' 4 'binge assessment just not reported' <p>(binges counted by episode - could be # of binges/day)</p>
Minimum binge freq (cut off) required for inclusion (e.g., 2x/wk for 6 months) (for diagnosis)			<ul style="list-style-type: none"> 1 'yes – minimum binge frequency' 2 'no – minimum binge frequency specifically not required' 3 'uncertain – dx assessed through interview/ questionnaire and binge frequency otherwise not mentioned' 4 'binge assessment just not reported'
Descriptive binge frequencies by			<ul style="list-style-type: none"> 1 'day' 2 'episode' 3 'uncertain – binges tracked descriptively through use of some tool (interview or questionnaire)' 3 'binge not described descriptively'
Lack of control over eating specifically required for binge			<ul style="list-style-type: none"> 1 'yes – lack of control specifically required' 2 'lack of control specifically not required' 3 'uncertain – dx was assessed through interview/questionnaire, lack of control otherwise not mentioned' 4 'lack of control assessment just not reported'
Distress regarding binge eating assessed			<ul style="list-style-type: none"> 1 'yes – specifically stated' 2 'distress specifically not required' 3 'uncertain – dx was assessed through interview/questionnaire, and distress otherwise not mentioned' 4 'distress assessment just not reported'
COMPENSATORY BX:			
Treatment of Compensatory Behaviors			<ul style="list-style-type: none"> 1 'irregular compensatory behaviors specifically allowed/included' 2 'compensatory behaviors specifically NOT allowed/ included' 3 'compensatory behaviors not mentioned as part of participant selection' 4 'compensatory behaviors present in BED group descriptively' (but not mentioned for participant selection)

Irregular compensatory behaviors frequency cut-off			1 'yes –participants were allowed to exhibit irreg. comp bxs at a cut-off frequency. List: 2 'yes – participants were allowed to exhibit irreg. comp bx, but no cut-off was reported' 3 'no –participants were not allowed to exhibit comp bxs at any frequency 3 'compensatory behavior not mentioned in any way'
Type of Compensatory behavior reportedly allowed in BED sample			1 'compensatory behaviors specifically not allowed at any frequency' 2 'only purging/vomit' 3 'only laxative / diuretic / enema abuse' 4 'only exercise' 5 'fasting/ restricting' 6 'multiple compensatory behaviors – List: ' 7 'compensatory behavior allowed, but type not reported' 8 'compensatory behavior not mentioned in any way'
<i>BODY WEIGHT:</i>			
Body weight assessed through			1 'BMI data reported descriptively 2 'American Insurance Industry's <i>Build & Blood Pressure Study</i> 3 'Metropolitan Life Company weight charts 4 'Lbs' 5 'Other' List:
Minimum BMI/Weight cut-off for participation			1 'yes' List: 2 'none reported'
Maximum BMI/Weight cut off for participation			1 'yes' List: 2 'none reported'
Obesity reported for some participants			1 'yes' 2 'no'
Actual time period of eating disturbance reported			1 'yes' 2 'no'

DUAL DIAGNOSIS			
BED as only Axis I or Axis II dx allowed?		s	1 'Yes - BED only was specifically stated' 2 'No - Dual or Mixed diagnoses specifically allowed' 3 'Unsure - not specifically stated'
Substance abuse specifically permitted		s	1 'Yes – substance abuse also present' 2 'No'
Personality d/os specifically permitted		s	1 'Yes – personality disorders also present' 2 'No'

Borderline Pers. DO specifically permitted		s	1 'Yes – borderline disorder also present' 2 'No'
Bipolar Pers DO specifically permitted		S	1 'Yes – bipolar personality disorder also present' 2 'No'
Mood/depressive disorders specifically permitted		s	1 'Yes - mood/depr disorders also present' 2 'No'
Anxiety disorders specifically permitted		s	1 'Yes - anxiety disorders also present' 2 'No'
Other diagnoses specifically permitted			1 'Yes - some other dx (besides above) permitted' (List here) 2 'No'
Obesity specifically permitted		s	1 'Yes - obesity also present' 2 'No'
Hx of other ED specifically permitted			1 'yes' 2 'no'
Medical problems assessed as part of the subject selection process			1 'yes' 2 'no' (ignore if medical problems followed only as a dependent or outcome variable)

EXCLUSIONARY CRITERIA			
Exclusionary criteria Mentioned		c	1 'Yes - exclusionary criteria (i.e., any other diagnosis) was mentioned or described' 2 'No'
Substance abuse specifically excluded?		s	1 'Yes - substance abuse excluded' 2 'No - substance abuse not excluded'
Personality disorders specifically excluded?		c	1 'Yes – personality disorders excluded' 2 'No – personality disorders not excluded'
Bipolar disorders specifically excluded?		s	1 'Yes - Bipolar disorders excluded' 2 'No - bipolar disorders not excluded'
Borderline disorder specifically excluded?		S	1 'Yes – Borderline disorder excluded' 2 'No – Borderline disorder not excluded'

Avoidant PD specifically excluded?		S	1 'Yes – avoidant p. disorder specifically excluded' 2 'No – avoidant personality disorder not excluded'
Mood/Depressive disorders specifically excluded?		s	1 'Yes - mood/depressive disorders excluded' 2 'No - mood/depressive disorders not excluded'
Anxiety Disorders specifically excluded?		s	1 'Yes - anxiety disorders excluded' 2 'No - anxiety disorders not excluded'
Psychotic disorders specifically excluded?		c	1 'Yes - psychotic disorders excluded' 2 'No - psychotic disorders not excluded'
Obese individuals specifically excluded?		c	1 'Yes - obese excluded' 2 'No - obese not excluded'
Hx of Other eating disorder specifically excluded			1 'Yes' List: 2 'No – history of other eating disorders not specifically excluded'
Zero evidence that Dual/Extra Diagnoses were considered?		c	1 'Yes - Dual diagnoses were simply not mentioned or referred to in any way' 2 'No - dual diagnoses were mentioned somehow (whether included or excluded)'
Other criteria specifically excluded			1 'Yes - some other criteria (besides above – e.g., no hx of prior med use/tx) excluded' (List:) 2 'No'

COMORBID HEALTH CONDITIONS			
List other medical problems tracked in patients with BN			List:

APPENDIX D

IRB WAIVER – NONHUMAN SUBJECTS STUDY



June 18, 2008

Rebecca Roberts
218 Lincoln Green Apt. F.
Starkville, MS 39759

Re: Dissertation Study
A Review of Diagnostic Methods in Most Cited Articles for Anorexia Nervosa,
Bulimia Nervosa, and Binge Eating Disorder

Dear Ms. Roberts:

IRB review is required for any project that constitutes human subject research as defined in 45 CFR 46. The relevant definitions are:

Human subject means a living individual about whom an investigator (whether professional or student) conducting research obtains (1) data through intervention or interaction with the individual, or (2) identifiable private information.

Research means a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge.

Based on the information provided in your email and in subsequent telephone correspondence, it appears that the above referenced study does not meet the human subject portion of the definition of human subject research as there is no intervention or interaction with individuals nor is there any identifiable private information. Therefore, the study does not constitute human subjects research and IRB review and approval is not necessary.

If I can provide any further assistance, please feel free to contact me.

Sincerely,

[For electronic submission]

Christine Williams
IRB Administrator

Office for Regulatory Compliance

210 Box 6225 • 70 Morgan Avenue • Beckman 3665 • Mississippi State, MS 39762 • (662) 325-3434 • FAX (662) 325-8775