306318643Y

Towards the Definition of an 'Eating Disorder'

Kristin Sonja Bohn

Lincoln College

Department of Psychiatry
Warneford Hospital
Oxford
OX3 7JX

Thesis submitted for the degree of Doctor of Philosophy
at the University of Oxford

Trinity Term 2006



Abstract

Towards the Definition of an 'Eating Disorder'

Kristin Sonja Bohn, Lincoln College, Oxford University, DPhil, Trinity Term 2006

This dissertation is concerned with the diagnosis and classification of clinical eating disorders such as anorexia nervosa and bulimia nervosa. There were four overarching aims. The first was to describe the clinical characteristics of patients with the neglected DSM-IV eating disorder diagnosis "eating disorder NOS" and compare them with those of patients with bulimia nervosa. It was found that the two groups were remarkably similar. The second was to examine how the classificatory problems associated with this diagnosis might be solved. Three solutions were proposed and the clinical utility of two of them examined. It was concluded that the best interim solution was to broaden the diagnostic criteria for anorexia nervosa and bulimia nervosa and re-label the remaining cases of eating disorder NOS as either binge eating disorder or as a new eating disorder diagnosis. The third aim was to derive an operational definition of what constitutes an "eating disorder" This involved developing an interview-based measure of functional impairment secondary to eating disorder features and administering it to a large sample of people exhibiting the full range of eating disorder psychopathology. Multivariate statistics identified specific severity levels on five eating disorder features that were strongly associated with the presence of a clinically significant level of impairment. These eating disorder features were: the pursuit of strict dietary rules, objective bulimic episodes, purging episodes, dissatisfaction with shape and weight, and over-concern with maintaining strict control over eating. The presence of two or more of these features above the identified thresholds was most predictive of a clinically significant level of impairment. Thus, an impairment-based, transdiagnostic, provisional operational definition of an eating disorder was derived. The fourth aim was to

develop a clinically useful, easily administered measure of psychosocial impairment secondary to eating disorder features. Such an instrument was created. Studies of its psychometric properties, reliability, validity and sensitivity to change all supported its use. Certain of the research strategies used in this dissertation could be usefully applied to other psychiatric disorders.

Acknowledgements

This thesis would not exist without the considerable support from Professor Christopher Fairburn and Dr Zafra Cooper. Professor Fairburn's unparalleled and infectious enthusiasm with regard to researching where the boundary lies between eating disorder and disordered eating inspired me to undertake this research and kept me motivated throughout. Dr Cooper's professional guidance and encouragement was invaluable. I am sincerely grateful to both of them for having taken on the supervision of my dissertation.

I would like to thank my colleagues and friends at the Centre for Research in the Eating Disorders (CREDO) in Oxford for having provided such a warm and inspiring environment to work in over the past five years. All of them have contributed to my research project in some way. I owe special thanks to Marianne O'Connor, Caroline Plumb, Elizabeth Payne, Shani Langdon and Jocasta Webb for having administered the CIA hundreds of times. Their feedback and ideas were crucial in the development of the instrument. I am also extremely grateful to Dr Helen Doll for her fantastic statistical guidance without which the completion of the main part of this dissertation would not have been possible. Special thanks to Mara Catling, Debbie Hawker, Roz Shafran, Michelle Lee and Rebecca Murphy for having been kind and encouraging. My colleagues at Leicester General Hospital have been equally invaluable. Professor Bob Palmer contributed substantially to the development of the CIA and Jackie Wales, Claire Nollett and Elizabeth Benson collected a large amount of data with it. I am very grateful for their ideas and hard work.

I would also like to express my thanks to all the women and men who participated in the study.

I feel immensely fortunate to have had the opportunity to undertake this research at Oxford. This would not have been possible without the financial support from the Wellcome Trust, which I gratefully acknowledge.

There are a number of people in my life outside work who have (indirectly) contributed immensely to this dissertation. My family back home in Germany, particularly my aunt Martina and my sister Kathrin, has been an important source of support and stability. My friends Dorle, Clare and Petra have been anchors in good and bad times. I would like them to know how grateful I am for their friendship. Very special thanks also to Gesa, Hauke & little sunshine Nora, and to Steffen, Catherine and Jocasta.

Finally I would like to express my gratitude to my partner Chris. I cannot say how much I have appreciated his support and encouragement. His enduring belief in me has been - and is - incredibly important. Thank you.

Table of Contents

CHAPTER ONE 9
1.1 Introduction 9
1.2 Chapter Overview 9
1.3 Diagnosis and Classification of Eating Disorders 10
1.3.1 Introduction 10
1.3.2 What is an eating disorder? 10
1.3.3 What is a mental disorder? 11
1.3.4 DSM and Eating Disorders 12
1.3.5 Shortcomings of the DSM-IV scheme 18
1.3.5.1 Criticisms of the diagnostic criteria for anorexia nervosa (AN) 18
1.3.5.2 Criticisms of the diagnostic criteria for bulimia nervosa (BN) 22
1.3.5.3 Criticism of the diagnosis eating disorder NOS 27
1.3.5.4 Criticisms of the DSM-IV scheme for classifying eating disorders 32
1.4 Eating Disorder Not Otherwise Specified (NOS) 38
1.4.1 Introduction 38
1.4.2 Clinical features of Eating Disorder NOS 39
1.4.3 Prevalence of Eating Disorder NOS 43
1.4.4 Course of Eating Disorder NOS 44
1.4.5 Treatment of Eating Disorder NOS 45
1.5 Clinical significance 47
1.5.1 Introduction 47
1.5.2 Impairment in the eating disorders 48
1.5.3 Conclusions and implications 55
1.6 Summary 56

CHAPTER TWO 57
2.1 General aims 57
2.2 Specific aims 57
2.3 Outline of the dissertation 59

CHAPTER THREE 61
3.1 Introduction 61
3.2 The Clinical Features of Eating Disorder NOS in Comparison with Bulimia Nervosa 62
3.2.1 Aim 62
3.2.2 Methods 62
3.2.3 Results 72
3.2.4 Summary of findings 81
3.3. Characterising Eating Disorder NOS further 82
3.3.1 Introduction and Aims 82
3.3.2 Methods 83
3.3.3 Results 85
3.3.3.1 Sub-analysis 90
3.3.4 Summary of findings 94
3.4 Discussion 95

CHAPTER FOUR 103
4.1 Introduction 103
4.2 Three solutions to the problem of Eating Disorder NOS 104
4.2.1 Relax the Diagnostic Criteria for AN and BN 104
4.2.2 Reclassify Eating Disorder NOS 106
4.2.3 The "Transdiagnostic" Solution 107
4.2.4 Summary 108
4.3 The impact of adjusting the DSM-IV criteria for AN and BN on the relative prevalence of Eating Disorder NOS 109
4.3.1 Introduction and aim 109
4.3.2 Methods 110
4.3.3 Results 113
4.3.4 Summary of findings 116
4.4 Evaluating the Clinical Utility of Fairburn and Bohn's (2005) "Interim Proposal" 117
4.4.1 Introduction 117
4.4.2 Aim 119
4.4.3 Methods 119
4.4.4 Results 123
4.4.5 Summary of findings 130
4.5 Discussion 132

CHAPTER FIVE 141
5.1 Introduction 141
5.2 Development of the Clinical Impairment Assessment Interview (CIA) 142
5.2.1 The content of the CIA 143
5.2.2 The rating scheme of the CIA 145
5.2.3 The style of the CIA 147
5.2.4 The use of the CIA 149
5.2.5 The CIA – from the preliminary to the current version 150
5.3 The Reliability and Validity of the CIA 155
5.3.1 Study One: Inter-rater Reliability of the CIA 157
5.3.2 Study Two: Test-retest Reliability of the CIA 160
5.3.3 Study Three: Construct and Discriminant Validity of the CIA 164
5.4 Identifying those eating disorder features that are associated with clinically significant secondary functional impairment 179
5.4.1 Introduction 179
5.4.2 Design 180
5.4.3 Participants 180
5.4.4 Measures 182
5.4.5 Data analysis 182
5.4.5.1 Pre-selection of suitable EDE items 182
5.4.5.2 Detailed analysis 184

5.4.5.3 Determination of cut-off points 186
5.4.6 Results 186
5.4.6.1 Pre-selection of suitable EDE items 186
5.4.6.2 Detailed analysis 193
5.4.6.3 Determination of the cut-off points on the final EDE items 199
5.4.7 Summary of findings 202
5.5 Discussion 203

CHAPTER SIX 213
6.1 Introduction 213
6.2 The Development of the CIA-Q 214
6.3 The Reliability and Validity of the CIA-Q 219
6.3.1 Study One: Internal consistency and factor structure of the CIA-Q 221
6.3.2 Study Two: Test-retest Reliability of the CIA-Q 228
6.3.3 Study Three: Construct and Discriminant Validity of the CIA-Q 231
6.3.4 Study Four: Sensitivity to Change of the CIA-Q 242
6.4 Discussion 245

CHAPTER SEVEN 252
7.1 Chapter overview 252
7.2 Restatement of aims and summary of findings 252
7.3 Conclusions of the research 254
7.4 Broader relevance of the research 255
7.5 Future directions 256

REFERENCES 258

INDEX OF TABLES 270

INDEX OF FIGURES 272

INDEX OF APPENDICES 273

APPENDICES 274

CHAPTER ONE

Introduction and Review of the Literature

1.1 Introduction

This dissertation is concerned with clinical eating disorders and their diagnosis and classification. Three eating disorders are currently recognised; anorexia nervosa, bulimia nervosa and variants of them (termed "eating disorder not otherwise specified" or "atypical eating disorders"). Together these disorders are a significant source of physical and psychosocial impairment among adolescent girls and young adult women: they are much less common among men. Three topics are central to the dissertation: first, the characteristics of patients with eating disorders other than anorexia nervosa and bulimia nervosa; second, the impairment that results from eating disorder psychopathology; and third, new ways of classifying eating disorders that are a better reflection of clinical reality than the current one. The dissertation opens with a review of the relevant literature and research. It is important to note, that the research in this dissertation focuses on the diagnosis and classification of eating disorders in adults. The implications of the research for the diagnosis and classification of eating disorders in children and adolescents are discussed in Chapter Seven.

1.2 Chapter Overview

In this chapter the literature relevant to the research is reviewed. Section 1.3 provides an overview of the diagnosis and classification of eating disorders in adults and highlights the problems with the main classificatory scheme. Section 1.4 reviews the literature on one specific eating disorder diagnosis, "eating disorder not otherwise specified (NOS)". This is the most common eating disorder diagnosis in most outpatient settings, yet it is the eating disorder that least is known about. In section 1.5 the topic of "clinical

significance" with regard to eating disorders is discussed. Studies of functional impairment in patients with eating disorders are also reviewed. Section 1.6 lists the main conclusions that may be drawn from this literature review.

1.3 Diagnosis and Classification of Eating Disorders

1.3.1 Introduction

The research in this dissertation focuses on the diagnosis and classification of eating disorders in adults. The implications of the research for the diagnosis and classification of eating disorders in children and adolescents are discussed in Chapter Seven.

1.3.2 What is an eating disorder?

The term 'eating disorder' is generally equated with the two syndromes 'Anorexia Nervosa' and 'Bulimia Nervosa'. Surprisingly, few attempts have been made to define an 'eating disorder' as such. One was coined by Fairburn and Walsh (2002), who suggested that:

> *"An eating disorder be defined as a persistent disturbance of eating behaviour or behaviour intended to control weight, which significantly impairs physical health or psychosocial functioning. This disturbance should not be secondary to any recognized general medical disorder (e.g., a hypothalamic tumour) or any other psychiatric disorder (e.g., an anxiety disorder)."*
>
> (Fairburn & Walsh, 2002; in Fairburn & Brownell, 2002, p.171)

This definition was later refined as follows:

"
- There is a definite disturbance of eating habits or weight control behaviour
- Either this disturbance, or associated core eating disorder features, results in a clinically significant impairment of physical health or psychosocial functioning (core eating disorder features comprise the disturbance of eating and any associated overevaluation of shape or weight)
- The behavioural disturbance should not be secondary to any general medical disorder or to any other psychiatric condition."

(Fairburn and Harrison, 2003, p.408)

Given the lack of an officially agreed upon definition of what constitutes an eating disorder, the work in this dissertation is based on the definition provided by Fairburn and Harrison (2003).

1.3.3 What is a mental disorder?

Two different schemes are currently used for classifying mental disorders. The first is the 'Diagnostic and Statistical Manual for Mental Disorders' (DSM) of the American Psychiatric Association (APA), and the second is the 'International Classification of Mental and Behavioural Disorders' (ICD) of the World Health Organisation (WHO). The APA's DSM is more widely used in eating disorder research than the ICD and thus the focus in this dissertation will be on the DSM classificatory scheme. The implications of the research for the current (tenth) version of the ICD (ICD-10; WHO, 1992) will be discussed in Chapter Seven. The current (fourth) edition of the DSM (DSM-IV; APA, 1994) conceptualises a mental disorder as a

> *"clinically significant behavioural or psychological syndrome or pattern that occurs in an individual and that is associated with present distress (...) or disability (...) or with a significantly increased risk of suffering death, pain, disability, or an important loss of freedom."*
>
> (APA, 1994, p.xxxi).

What makes a behavioural or psychological phenomenon "clinically significant" is clearly a critical issue when considering the degree of disturbance that has to be present for a state to be viewed as a mental disorder. In this context the following is stated in DSM-IV:

> *"The definition of mental disorder (...) requires that there be clinically significant impairment or distress. To highlight the importance of considering this issue, the criteria sets for most disorders include a clinical significance criterion (usually worded "...causes clinically significant distress or impairment in social, occupational, or other important areas of functioning"). This criterion helps establish the threshold for the diagnosis of a disorder in those situations in which the symptomatic presentation by itself (particularly in its milder forms) is not inherently pathological and may be encountered in individuals for whom a diagnosis of "mental disorder" would be inappropriate. Assessing whether this criterion is met, especially in terms of role function, is an inherently difficult clinical judgment."*

(APA, 1994, p.7)

This statement indicates that there is no standardised definition of "clinical significance" and that it is essentially a clinical judgment.

1.3.4 DSM and Eating Disorders

In DSM-IV two specific eating disorders are recognised; anorexia nervosa and bulimia nervosa. Both diagnoses have a set of diagnostic criteria that need to be fulfilled for the diagnoses to be made (see below). In addition, there is a third diagnostic category, "Eating Disorder Not Otherwise Specified" (eating disorder NOS). This "NOS" appellation requires some explanation. In DSM-IV each class of mental disorder contains one or more specific categories, each of which is defined by a set of diagnostic criteria. In addition, each class contains one 'not otherwise specified (NOS)' category. The latter is a residual category meant to cover states that are of clinical severity (i.e., cause clinically significant impairment or distress, as noted above) but do not conform to the diagnostic criteria of the specific categories (i.e., in this case anorexia nervosa or bulimia nervosa). The NOS category "indicates a category within a class of disorders that is residual to the specific categories in that class (...)" (APA, 1980, p.32; APA, 1987, p.23). The main rationale for the existence of NOS diagnostic categories is, according to DSM-IV, "the diversity of clinical presentations" that makes it "impossible for the diagnostic nomenclature to cover every possible situation" (APA, 1994, p.4). The implication of this statement and the term "residual" is that NOS diagnoses should be uncommon.

Anorexia Nervosa

Table 1.1 shows the DSM-IV diagnostic criteria for anorexia nervosa. Criterion A requires that people with this diagnosis actively maintain an unusually low body weight. Precisely what constitutes being below "a minimally normal weight for age and height" is

left to the clinician's judgment, but DSM-IV provides, as a suggested guideline, weighing less than 85% of that expected. Since this is rather vague, many clinicians adopt the ICD-10 threshold: 15% below the expected weight for age and height, i.e., a body mass index (weight in kilograms divided by the height in meters squared) equal to or less than 17.5 kg/m^2 (Turner & Bryant-Waugh, 2004; Treasure, 1999). Criterion B requires the presence of an intense fear of weight gain or fatness, even though underweight. Criterion C requires that a person with anorexia nervosa evaluates themselves almost exclusively on the basis of their body weight and/or shape. This is often reflected in repeated body checking and feelings of fatness. Criterion D requires that women with anorexia nervosa do not menstruate. The diagnosis of anorexia nervosa also contains two mutually exclusive subtypes; the restricting and the binge/purging type.

Table 1.1 *DSM-IV diagnostic criteria for Anorexia Nervosa (APA, 1994, p.544)*

DSM-IV Diagnostic Criteria for Anorexia Nervosa	
A.	Refusal to maintain one's body weight at or above a minimally normal weight for age and height (e.g., weight loss leading to maintenance of body weight less than 85% of that expected; or failure to make expected weight gain during period of growth, leading to body weight less than 85% of that expected).
B.	Intense fear of gaining weight or becoming fat, even though underweight.
C.	Disturbance in the way in one's body weight or shape is experienced, undue influence of body weight or shape on self-evaluation, or denial of the seriousness of the current low body weight.
D.	In postmenarcheal females, amenorrhea, i.e., the absence of at least three consecutive menstrual cycles.
Restricting type: During the current episode of anorexia nervosa, the person has not regularly engaged in binge eating or purging behavior (i.e., self-induced vomiting or the misuse of laxatives, diuretics, or enemas).	
Binge Eating/Purging Type: During the current episode of anorexia nervosa, the person has regularly engaged in binge eating or purging behavior (i.e., self-induced vomiting or the misuse of laxatives, diuretics, or enemas).	

Bulimia Nervosa

The DSM-IV diagnostic criteria for bulimia nervosa are shown in Table 1.2. Bulimia nervosa is characterised by recurrent episodes of binge eating (criterion A). The binges have to be large in size and have to have occurred within a two-hour period. In addition, the individual must have had experienced a sense of lack of control during the binge. A further defining criterion of bulimia nervosa is the recurrent use of compensatory weight-control behaviour (criterion B). As in anorexia nervosa, over-evaluation with weight and shape is also required (criterion D). Criterion C gives a severity threshold for the binge eating and the compensatory behaviour. Criterion E clarifies the cross-sectional relationship between anorexia nervosa and bulimia nervosa by specifying that if the criteria for both disorders are fulfilled the diagnosis of anorexia nervosa is made. Bulimia nervosa, too, is divided into two subtypes, a purging and non-purging type.

Table 1.2 *DSM-IV diagnostic criteria for Bulimic Nervosa (APA, 1994, p.549)*

DSM-IV Diagnostic Criteria for Bulimia Nervosa	
A.	Recurrent episodes of binge eating. An episode of binge eating is characterized by both of the following: (1) eating, in a discrete period of time (e.g., within any 2-hour period), an amount of food that is definitely larger than most people would eat during a similar period of time and under similar circumstances (2) a sense of lack of control over eating during the episode (e.g., a feeling that one cannot stop eating or control how much one is eating).
B.	Recurrent inappropriate compensatory behavior in order to prevent weight gain, such as self-induced vomiting; misuse of laxatives, diuretics, enemas, or other medications; fasting; or excessive exercise.
C.	The binge eating and inappropriate compensatory behaviors both occur, on average, at least twice a week for 3 months.
D.	Self-evaluation is unduly influenced by body shape and weight.
E.	The disturbance does not occur exclusively during episodes of Anorexia Nervosa.

Purging type: During the current episode of bulimia nervosa, the person has regularly engaged in self-induced vomiting or the misuse of laxatives, diuretics, or enemas.

Nonpurging type: During the current episode of bulimia nervosa, the person has used other inappropriate compensatory behaviors, such as fasting or excessive exercise, but has not regularly engaged in self-induced vomiting or the misuse of laxatives, diuretics, or enemas.

Eating Disorder Not Otherwise Specified (Eating Disorder NOS)

The third eating disorder category, eating disorder NOS, is an example of the residual 'NOS' category as noted above. It was first introduced as "Atypical Eating Disorder" in the third edition of the DSM (APA, 1980). It is reserved for coding "disorders of eating that do not meet the criteria for any specific Eating Disorder" (APA, 1994, p.550). It therefore denotes those conditions that meet the definition of an eating disorder but not the criteria for anorexia nervosa or bulimia nervosa. In contrast to the other two eating disorders, it does not have any positive diagnostic criteria other than the requirement that there be an eating disorder of clinical severity. Therefore, two steps are involved in making a diagnosis of eating disorder NOS (Fairburn and Bohn, 2005):

1. It needs to be determined whether an eating disorder of clinical severity is present,
2. It then needs to be established that the diagnostic criteria for anorexia nervosa and bulimia are not met.

The diagnosis is therefore made by exclusion. How to determine the presence of the only positive criterion for the diagnosis (viz. an 'eating disorder of clinical severity') is left to the judgment of the clinician.

It cannot be stressed too strongly that to make a diagnosis of eating disorder NOS the person must have an eating disorder of clinical severity. This means that the eating disorder features need to result in clinically significantly impairment in psychosocial or physical functioning. This point has been neglected in studies of non-clinical samples (Fairburn and Bohn, 2005).

In an attempt to provide clinicians with guidance as to how to make a diagnosis of eating disorder NOS, DSM-IV provides a number of examples of the types of clinical state that would be classified as eating disorder NOS (APA, 1994, p.550). These are shown in Table 1.3.

Table 1.3 *DSM-IV examples of Eating disorder NOS (APA, 1994, p.550)*

	DSM-IV Examples of Eating disorder NOS
1.	For females, all of the criteria of anorexia nervosa are met except that the individual has regular menses.
2.	All of the criteria for anorexia nervosa are met except that, despite significant weight loss, the individual's current weight is in the normal range.
3.	All of the criteria for BN are met except that the binge eating and inappropriate compensatory mechanisms occur at a frequency of less than twice a week or for a duration of less than three months.
4.	The regular use of inappropriate compensatory behavior by an individual of normal body weight after eating small amounts of food (e.g., self-induced vomiting after the consumption of two cookies).
5.	Repeatedly chewing and spitting out, but not swallowing, large amounts of food.
6.	Binge-eating disorder: recurrent episodes of binge eating in the absence of the regular use of inappropriate compensatory behaviors characteristic of BN.

Examples 2 to 4 in Table 1.3 describe clinical states that might best be described as subthreshold forms of anorexia nervosa and bulimia nervosa. They resemble the specified eating disorder, but do not meet its exact diagnostic criteria, because the frequency or severity of one or more features is below the specified threshold. In contrast, examples 1, 5 and 6 might better be viewed as atypical (i.e., they show a clinical picture that is different from anorexia nervosa and bulimia nervosa yet is characterised by eating disorder features of clinical significance).

A recent development of relevance is the proposal that a third 'specified' eating disorder be recognised to accompany anorexia nervosa and bulimia nervosa. This new diagnosis has been termed "binge eating disorder". It is intended for people who experience

recurrent episodes of binge eating in the absence of the extreme methods of weight control seen in bulimia nervosa and anorexia nervosa (APA, 1994). This proposal was controversial when it was first suggested (Fairburn, Welch, & Hay, 1993; Spitzer, Stunkard, Yanovski, Marcus, Wadden et al., 1993) and divergent views on its merit still exist (Stunkard & Allison, 2003; Wilfley, Wilson, & Agras, 2003). Provisional research criteria for binge eating disorder have been provided in Appendix B of DSM-IV ("Criteria Sets and Axes Provided for Further Study", p.731) in order to facilitate research on the topic. They are shown in Table 1.4. As matters stand binge eating disorder is not an established DSM-IV diagnosis and therefore eating disorders of this type remain a form of eating disorder NOS (see Table 1.3). It is for this reason that they will not be considered as a separate diagnosis in this dissertation.

Table 1.4 *DSM-IV research criteria for Binge eating disorder (BED) (APA, 1994, p.731)*

DSM-IV Research Criteria for Binge-eating disorder	
A.	Recurrent episodes of binge eating. An episode of binge eating is characterized by both of the following: (1) eating, in a discrete period of time (e.g., within any 2-hour period), an amount of food that is definitely larger than most people would eat during a similar period of time and under similar circumstances (2) a sense of lack of control over eating during the episode (e.g., a feeling that one cannot stop eating or control how much one is eating).
B.	The binge-eating episodes are associated with three (or more) of the following: (1) eating much more rapidly than normal (2) eating until feeling uncomfortably full (3) eating large amounts of food when not feeling physically hungry (4) eating alone because of being embarrassed by how much one is eating (5) feeling disgusted with oneself, depressed, or very guilty after overeating
C.	Marked distress regarding binge eating is present.
D.	The binge eating occurs, on average, at least 2 days a week for 6 months. **Note:** The method of determining frequency differs from that used for Bulimia Nervosa; future research should address whether the preferred method of setting a frequency threshold is counting the number of days on which binges occur or counting the number of episodes of binge eating.
E.	The binge eating is not associated with the regular use of inappropriate compensatory behaviors (e.g., purging, fasting, excessive exercise) and does not occur exclusively during the course of Anorexia Nervosa or Bulimia Nervosa.

1.3.5 Shortcomings of the DSM-1V scheme

The DSM-IV scheme for classifying and diagnosing eating disorders has been the subject of much debate. In this section, the main criticisms of the overall scheme are reviewed together with those of the specific diagnostic criteria for the three eating disorders.

1.3.5.1 Criticisms of the diagnostic criteria for anorexia nervosa

In this section each of the diagnostic criteria for anorexia nervosa will be discussed in turn.

Criterion A: The weight threshold

The cardinal feature of anorexia nervosa is that patients with this disorder are significantly underweight. Where exactly the threshold for "significantly underweight" should be drawn has been debated extensively by clinicians and researchers in the field of eating disorder research. There has also been some variation in the definition of the weight threshold for anorexia nervosa over the years. In the third edition of DSM (DSM-III; APA, 1980) the weight criterion was "weight loss to at least 25% of original body weight". It was then changed to "maintaining a body weight 15% below expected" in DSM-III-R (APA, 1987). In DSM-IV the criterion is "weight loss leading to the maintenance of a body weight below 85% expected" (APA, 1994). This variation illustrates the uncertainties pertaining to the specification of the cut-off point to be used.

The main criticism with regard to the current cut-off is that it is arbitrary (i.e., not based on empirical data) (e.g., Mitchell, Cook-Myers, & Wonderlich, 2005; Watson & Andersen, 2003; Hebebrand, Casper, Treasure, & Schweiger, 2004). Few studies have determined at what weight the symptoms of starvation that are typical of anorexia nervosa actually start to appear or at what weight the adverse physiological consequences of a low weight become evident (Garfinkel, Kennedy, & Kaplan 1995a). In a recent study by Watson

and Andersen (2003) 297 patients with DSM-IV anorexia nervosa were compared with 67 patients with eating disorder NOS on numerous variables ranging from demographic characteristics, history of the eating disorder, level of depression, weight history, and eating disorder psychopathology at the time point of admission to an in-patient clinic. The patients with eating disorder NOS fulfilled all diagnostic criteria for anorexia nervosa with the exception of the weight criterion and the amenorrhoea criterion. Few differences between the two groups were found. The authors concluded that "modifying the DSM-IV weight criterion (...) for anorexia nervosa will be more beneficial than harmful" (Watson & Andersen, 2003) and they suggested the weight criterion for anorexia nervosa could be changed to requiring a certain amount of weight loss, rather than the attainment of a specific weight.

Other researchers have suggested that the weight threshold for anorexia nervosa in DSM-IV is too strict and should be adjusted upwards (e.g., Thaw, Williamson, & Martin, 2001). Thaw and colleagues (2001) systematically modified two of the DSM-IV criteria of anorexia nervosa (they adjusted the weight threshold upwards from 15% to 10% underweight and deleted the amenorrhoea criterion) and examined the impact on the base rates of all three eating disorders. They found that their modifications, in isolation and combined, only had a minor impact on the relative prevalence of the three eating disorders. The clinical relevance of this finding is questionable, however, since Thaw and colleagues used a convenience sample of eating disorder cases, not a clinical one.

Criteria B and C: Fear of gaining weight and Over-evaluation of shape and weight

These two criteria have been criticised because of research showing that there are a substantial number of underweight patients who do not show an intense fear of gaining weight or becoming fat or do not over-evaluate shape and weight, but otherwise fulfil the

diagnostic criteria for anorexia nervosa. These patients are found in two main groups, non-Western patient samples and adolescent patients with a short history (Fairburn & Gowers, in press). Lee, Ho and Hsu (1993) reported that among anorexia nervosa patients in Hong Kong, concerns about weight are frequently not present yet the patients appear to have essentially the same disorder. Consequently, Lee and colleagues argued that weight concerns should not be an essential criterion of anorexia nervosa. There have also been descriptions of Western patients lacking a 'drive for thinness'. Ramacciotti and colleagues (Ramacciotti, Dell'Osso, Paoli, Ciapparelli, Coli et al., 2002) studied a group of anorexia nervosa patients who had been referred to a day-hospital in Canada. They compared patients with a low 'drive for thinness'[1] with those with a high drive for thinness. They found that 17% of the total sample lacked the drive for thinness yet the two groups were remarkably similar in other respects. Interestingly, Strober, Freeman and Morrell (1999) found a similar proportion (21%) of 'atypical anorexia nervosa' patients (lacking fear of weight gain and body size distortion) among a sample of inpatients with anorexia nervosa in the USA.

Palmer (2003) has stated that patients lacking weight or shape concerns could be encompassed into the diagnosis of anorexia nervosa by making a modest change to the diagnostic criteria. He suggested that "motivated eating restraint" be a core component of the disorder, instead of the current focus on shape and weight concerns (which can be viewed as one motivation for such eating restraint). A similar suggestion has been to redefine the 'core psychopathology' of anorexia nervosa (i.e., the over-evaluation of weight and shape) to include states in which there is an over-evaluation of controlling eating per se without requiring that there also be accompanying concerns about shape and weight (e.g., Rieger, Touyz, Swain, & Beumont, 2001). Reflecting this viewpoint, the latest edition of the

[1] A low drive for thinness was here defined as a rating of 7 or below on the subscale 'Drive for thinness' on the Eating Disorder Inventory-II (EDI-II; Garner, Olmsted & Polivy, 1983).

Eating Disorder Examination assessment interview (EDE, version 15; Fairburn & Cooper, 1993) has an item "Importance of Maintaining Strict Control over Eating" in addition to the "Importance of Shape" and "Importance of Weight" items.

Criterion D: Amenorrhoea

This is the most controversial diagnostic criterion of anorexia nervosa. Numerous clinicians and researchers have argued that the absence of menses should not be necessary for a diagnosis of anorexia nervosa to be made (Cachelin & Maher, 1998; Garfinkel et al., 1995a; Garfinkel, Lin, Goering, Spegg, Goldbloom et al., 1996; Watson & Andersen, 2003; Mitchell et al., 2005; Abraham, Pettigrew, Boyd, Russell, Taylor, 2005).

Amenorrhoea was initially introduced as a diagnostic criterion for anorexia nervosa because it was believed that the disorder might be due to primary impairment of hypothalamic functioning (Russell, 1969). Since then, however, it has been observed that amenorrhoea generally follows the onset of weight loss and that it appears to be a direct consequence of the loss of weight or percent body fat (Garfinkel & Garner, 1982). In addition, amenorrhoea does not only occur in patients with anorexia nervosa, but has also been observed in normal weight bulimia nervosa patients (Fairburn & Cooper, 1984). In a study investigating the question whether amenorrhoea is a useful criterion for anorexia nervosa, Cachelin and Maher (1998) compared twelve non-amenorrheic women treated for anorexia nervosa with 40 women meeting full DSM-IV criteria. The non-amenorrheic group displayed the same high levels of eating disorder features, body-image disturbance, and psychopathology as the amenorrheic one. The authors concluded that amenorrhoea is not a useful criterion for diagnosing anorexia nervosa. Similarly, Garfinkel and colleagues (1996) compared 24 women with a lifetime diagnosis of DSM-III-R anorexia nervosa to 44 women fulfilling all DSM-III-R diagnostic criteria for anorexia nervosa with the exception of the

amenorrhoea criterion. They found no statistically significant differences between the two groups with regard to their demographic features, current comorbid psychiatric features, the frequency of bulimic behaviour and parental psychiatric features. The authors concluded that "there does not appear to be value in defining a case by the presence of amenorrhoea" (Garfinkel et al, 1996).

1.3.5.2 Criticisms of the diagnostic criteria for bulimia nervosa

Criterion A: Recurrent episodes of binge eating

The presence of binge eating is a core criterion for the diagnosis of bulimia nervosa. However, the exact definition of a binge has been, and continues to be, the subject of debate and investigation. One element of the definition that has been criticised is the requirement that the amount of food eaten be large in size. Empirical research on the size of typical binges indicates that there is a great variability between episodes even within a single individual (Rossiter & Agras, 1990), and that there are many people with eating disorders who do not eat a large amount of food in their "binges" (Walsh, 1993). This is recognised in the EDE interview where "objective" and "subjective" bulimic episodes are distinguished, the former being episodes in which an unambiguously large amount of food is eaten (given the circumstances), and the latter being episodes in which a smaller amount is consumed (Fairburn & Cooper, 1993) (see Table 1.5).

Table 1.5 *Distinction between objective bulimic episodes (OBE) and subjective bulimic episodes (SBE) as defined in the Eating Disorder Examination (Fairburn & Cooper, 1993)*

		Loss of Control?	
		YES	**NO**
Large Amount of Food?	**YES**	OBE	Objective Overeating
	NO	SBE	Subjective Overeating

Researchers and clinicians have questioned the clinical usefulness of the DSM-IV requirement that binges be large in size (i.e., objective bulimic episodes) and have suggested that it be removed (e.g., Le Grange, Binford, Peterson, Crow, Crosby et al., 2006; Pratt, Niego & Agras, 1998; Rossiter & Agras, 1990). Studies indicate that there are no meaningful differences in psychopathology between individuals with objective binges compared with those with subjective ones. For example, Le Grange and colleagues (2006) compared 138 women with DSM-IV bulimia nervosa with 57 women with eating disorder NOS. The sample was a mixture of clinical and community cases. The eating disorder NOS cases fulfilled all diagnostic criteria for bulimia nervosa with the exception of one of the following two: either their binges were small in size (N=23) or their frequency was below the twice-weekly threshold but at least once per week (N=34). With the exceptions of ethnicity (more diverse in bulimia nervosa subjects), medication use, history of anorexia nervosa and eating concerns (all higher in the bulimia nervosa subjects), the women with bulimia nervosa and those with eating disorder NOS were not statistically different. The authors concluded that "these results should prompt re-evaluation of existing bulimia nervosa diagnostic boundaries (e.g., frequency criteria reduced from two to one bulimic episode per week, or eliminating the requirement that binge eating episodes involve the consumption of objectively large amounts of food)" (Le Grange et al., 2006).

In a similar study Pratt and colleagues (1998) assessed 174 women seeking treatment for DSM-IV bulimia nervosa. The authors created two measures of binge eating, namely the sum of objective and subjective episodes and the difference between number of objective and subjective ones. Although total number of episodes was associated with a measure of self-efficacy, the relative proportion of subjective episodes was not associated with any measure of severity of psychopathology. The authors concluded that the consumption of an

objectively large amount of food might not be a valid criterion for defining binge eating. It is important to note that all the women in this study met DSM-IV criteria for bulimia nervosa and therefore all experienced objective episodes at least twice per week over a three month period. This common feature might account for the failure to find meaningful differences within this subject group.

Overall these studies suggest that individuals who do not consume a large amount of food during their binges but otherwise fulfil the diagnostic criteria for bulimia nervosa are not different to those with DSM-IV bulimia nervosa with regard to their eating disorder and general psychiatric psychopathology. However, since there are only few studies to date that have examined the size of binges, the evidence is weak and needs to be supported by further studies.

Criterion B: Recurrent inappropriate compensatory behaviour

This criterion is another defining feature of bulimia nervosa but has not been criticised as such. There have been criticisms however of the types of compensatory behaviours specified in DSM-IV. More specifically, the compensatory behaviour 'fasting' and the exact definition of 'excessive exercise' have been the subject of debate (e.g., Cooper & Fairburn, 2003; Mond, Hay, Rodgers, Owen, & Beumont, 2004; Mond, Hay, Rodgers, & Owen, 2006). These are the two compensatory behaviours that are required to make the diagnosis of non-purging bulimia nervosa (see Table 1.2 above).

The 'fasting' criterion for non-purging bulimia nervosa was introduced in DSM-IV. In the previous version of the DSM (DSM-III-R; APA, 1987) "sustained and marked dietary restriction" was required. The reason for this change was the attempt to better delineate cases of binge eating disorder from those of non-purging bulimia nervosa. Cooper and Fairburn (2003) have pointed out that this restrictive definition of non-purging bulimia

nervosa has resulted in the "virtual elimination" of the diagnosis since very few people engage in the two compensatory behaviours as currently defined. They suggest that the diagnostic criteria for bulimia nervosa should be changed by removing the DSM-IV "fasting" requirement and replacing it with "sustained and marked dietary restriction" as before.

There has also been a debate about the definition of "excessive exercise" (e.g., Mond et al., 2004; Mond et al., 2006). DSM-IV states that "exercise may be considered to be excessive when it significantly interferes with important activities, when it occurs at inappropriate times or in inappropriate settings, or when the individual continues to exercise despite injury or other medical complications" (APA, 1994, p. 546). This definition leaves considerable room for individual variation in diagnostic practice, especially with regard to the required frequency of the behaviour. In an attempt to specify operational criteria for excessive exercise, Mond and colleagues (2004) examined relationships between exercise behaviour, eating-disordered behaviour and quality of life in a community sample of Australian women. The authors found that obligatory exercise (i.e., feeling guilty following the postponement of exercise) and exercising for shape and weight reasons (i.e., to improve physical appearance or body tone) were associated with higher levels of eating disorder psychopathology and reduced quality of life, while there was no association between exercise behaviour and quality of life after controlling for the effects of eating disorder psychopathology. They suggest that operational definitions of "excessive exercise" might usefully include reference to these variables. Both, obligatory exercise and exercising for shape and weight reasons have been incorporated in the definition of "Driven Exercising" in the Eating Disorder Examination 15.0 (EDE; Fairburn & Cooper, 1993; see Appendix 3.1), which is used in Chapter Four of this dissertation.

While this study contributes to knowledge about the quality of exercise that is excessive or driven, it does not shed any light on the quantity of this type of exercise that is required for it to be regarded as excessive. Further studies are needed to clarify the definition of this compensatory behaviour.

Criterion C: Frequency requirement for binge eating and compensatory behaviours (twice per week for three months)

This frequency requirement for binge eating was introduced in DSM-III-R (APA, 1987) due to concerns about the potential for over-diagnosing bulimia nervosa in the context of reports of an epidemic of bulimia nervosa (Johnson, Stuckey, Lewis & Schwartz, 1983). Adding it reduced prevalence estimates ten-fold in some cases (Ben-Tovim, 1988). However, since its introduction there have been concerns about the arbitrariness and clinical validity of the specified frequency threshold for both the binge eating and the compensatory behaviour (e.g., Garfinkel, Lin, Goering, Spegg, Goldbloom et al., 1995b; Wilson & Eldredge, 1991). While they serve the purpose of restricting the diagnosis to those individuals who have a recurrent and persistent problem, they may have the unfortunate consequence of excluding individuals whose problems are very similar to those of people with bulimia nervosa.

Garfinkel and colleagues (1995b) conducted a community study in which they compared 62 women with DSM-III-R bulimia nervosa with 22 women who were missing only the frequency criterion (two or more binge eating episodes per week for three months) but otherwise fulfilled DSM-III-R criteria for bulimia nervosa. No statistically significant differences between the two groups were found with regard to demographic characteristics, eating disorder features (with the exception of binge eating), comorbid psychiatric disorders, history of childhood sexual abuse and parental psychopathology. The authors concluded that the findings suggested "that while the frequency criterion continues to identify a

disturbed group with regard to comorbidities and environmental difficulties, it remains an arbitrary threshold that excludes from diagnosis subjects who in every other way resemble women with bulimia nervosa" (Garfinkel et al., 1995b, p.1057).

Wilson and Eldredge (1991) examined the concurrent and predictive validity of the twice weekly frequency thresholds for binge eating and compensatory behaviours. They classified treatment-seeking eating disorder patients according to different binge eating and purging frequencies and found that the frequency of purging (i.e., self-induced vomiting and laxative misuse), but not binge eating, was related to eating disorder psychopathology and general psychopathology. Neither purging nor binge eating frequencies predicted post-treatment outcome. The findings of this study need to be interpreted with caution, however, given the small sample size of the study (N=30).

These and other studies (e.g., Le Grange et al., 2006) suggest that the current twice weekly threshold for binge eating is not a clinically valid one in that it does not separate qualitatively different groups of patients with an eating disorder. For this reason treatment studies are now beginning to use a broader frequency threshold (e.g., Walsh, Fairburn, Mickley, Sysko & Parides, 2004). This said, more data needs to be collected on the eating disorder features, general psychopathology and treatment response of individuals who present with binge eating and compensatory behaviours at different frequencies.

1.3.5.3 Criticism of the diagnosis eating disorder NOS

In contrast to the two specific eating disorders anorexia nervosa and bulimia nervosa, eating disorder NOS does not have any positive diagnostic criteria apart from the requirement that states classified as it are eating disorders of clinical severity. Therefore the criticisms directed at the diagnosis are focused on the absence of any specific diagnostic criteria and the anomalous status of the diagnosis.

The literature concerning the diagnosis eating disorder NOS was synthesised by the candidate and Professor Christopher Fairburn in a review published in *Behaviour Research and Therapy* in 2005 (Fairburn and Bohn, 2005; see Appendix 1.1). The main points of relevance made in this review were as follows.

1. A problem of nosology

The diagnostic category eating disorder NOS is the DSM-IV 'residual' NOS category within the diagnostic class "Eating disorders" This literally means that it is a category for what is 'left over' in the pool of individuals with eating disorders of clinical severity, once those cases that fulfil diagnostic criteria for anorexia nervosa and bulimia nervosa have been taken out. As discussed earlier, NOS diagnoses are supposed to be infrequent comprising only few individuals.

This is not the case with regard to the diagnosis eating disorder NOS. In fact the opposite is the case: eating disorder NOS is the most common eating disorder diagnosis made in most adult outpatient settings other than those that attract highly specialist referrals. Table 1.6 shows the prevalence figures from five well-diagnosed adult samples. In each study eating disorder NOS was the most common diagnosis made, the weighted average prevalence across the five samples being 54.2%. It is important to note that the high proportion of eating disorder NOS cases in these samples was not due to laxity in defining what is an "eating disorder of clinical severity" (i.e., a "case") for the data were from people seeking treatment in whom an eating disorder diagnosis had been substantiated by a clinician.

Table 1.6 *Prevalence of eating disorder NOS in samples of adult outpatients with eating disorders*

	Sample size	DSM-IV diagnosis			Comments
		Anorexia nervosa %	**Bulimia nervosa %**	**Eating disorder NOS %**	
Millar et al. (1998)	510	13.5	39.6	46.9	Sample excluded patients with BED
Martin et al. (2000)	175	19.4	22.9	57.7	BED[1] 9.7% of full sample
Ricca et al. (2001)	189	24.9	24.9	50.3	BED 8.5% of full sample
Turner & Bryant-Waugh (2004)	190	5.8	23.7	70.5	Sample excluded patients with BED
Nollett & Button (2005)	187	16.0	25.7	58.3	BED 8.0% of full sample
Weighted average		**15.3**	**30.5**	**54.2**	

[1] BED - Binge eating disorder

The high prevalence of the diagnosis eating disorder NOS in adult outpatient samples is clearly problematic. The residual category should not be the most common one. The fact that it is, suggests that there is something fundamentally wrong with the scheme for classifying eating disorders.

2. The absence of positive diagnostic criteria

The absence of an agreed definition of what constitutes eating disorder NOS leaves considerable room for individual variation in diagnostic practice. This situation is quite different to that existing for anorexia nervosa and bulimia nervosa where a specific combination of clinical features must be present for either diagnosis to be made. It is helpful to illustrate diagrammatically the cross-sectional relationship between the diagnoses anorexia nervosa, bulimia nervosa and eating disorder NOS (see Figure 1.1). The two overlapping inner circles represent anorexia nervosa (the smaller circle) and bulimia nervosa

(the larger circle) respectively, the area of potential overlap being that occupied by those people who would meet the diagnostic criteria for both disorders but for the DSM-IV "trumping" rule whereby the diagnosis of anorexia nervosa takes precedence over that of bulimia nervosa. Surrounding these two circles is an outer circle which defines the boundary of eating disorder "caseness"; that is, the boundary between having an eating disorder, a state of clinical significance, and having a lesser, non-clinical, problem with eating. It is this boundary that demarcates what is, and is not, an eating disorder. Within the outer circle, but outside the two inner circles, lies eating disorder NOS.

Figure 1.1 *A schematic representation of the relationship between anorexia nervosa, bulimia nervosa and eating disorder NOS (taken from Fairburn & Bohn, 2005)*



Given this diagnostic scheme, the challenge involved in formulating diagnostic criteria for eating disorder NOS lies in defining its outer "edges" (as illustrated in Figure 1.1) since the inner boundaries, those of anorexia nervosa and bulimia nervosa, are already defined (although suggestions have been made for how to adjust them, as described above). It is one of the goals of the research described in this dissertation to propose a definition of

the "outer edges" of eating disorder NOS and therefore the boundary of eating disorder "caseness" (see Chapter Five).

3. The neglect of the diagnosis eating disorder NOS

Most of the research on eating disorders has focused on the two specified eating disorder diagnoses anorexia nervosa and bulimia nervosa. Over the last decade there has been somewhat more interest in eating disorders other than anorexia nervosa and bulimia nervosa but so far few studies have focused on eating disorder NOS as a whole (Button, Benson, Nollett & Palmer, 2005; Mizes & Sloan, 1998; Ricca, Mannucci, Mezzani, Di Bernardo, Zucchi, Paionni et al., 2001; Turner & Bryant-Waugh, 2004). Most studies investigating eating disorders other than anorexia nervosa or bulimia nervosa have merely broadened the diagnostic criteria for the two specified diagnoses to include some 'subthreshold' or 'atypical' cases (e.g., Cotrufo, Monteleone, Castaldo & Maj, 2004; Crow, Agras, Halmi, Mitchell & Kraemer, 2002; Herzog, Hopkins, & Burns, 1993), but by doing so they have not studied a substantial number of eating disorder NOS cases (see Chapter Four). No studies of the treatment of eating disorder NOS have been reported at the current time.

A likely reason for this neglect of the most common eating disorder diagnosis is its 'not otherwise specified' status. NOS diagnoses in general are not much studied (Pincus, Wakefield Davies, & McQueen, 1999). In some countries this neglect of eating disorder NOS has a direct impact on patient care for the marginal status of NOS diagnoses extends even to restrictions on treatment provision or, at least, reimbursement for treatment (Herzog et al., 1993; Martin, Williamson, & Thaw, 2000). This could perhaps be justified were NOS states uncommon but this could not be said to be true of eating disorder NOS.

Another potential reason for the neglect of this diagnosis, and patients who receive it, is the common misconception that cases of eating disorder NOS do not have a true eating disorder but have a milder "subclinical" eating disorder. This view is clearly inappropriate given that the only "positive" criterion for eating disorder NOS is that there needs to be an eating disorder of clinical severity. Furthermore, as described below, there is now evidence that the severity of psychopathology and degree of secondary psychosocial impairment in those with eating disorder NOS are comparable to those seen in patients with anorexia nervosa and bulimia nervosa (Ricca et al., 2001; Turner & Bryant-Waugh, 2004; see Chapter Three).

A positive step towards recognising eating disorders other than anorexia nervosa and bulimia nervosa was the introduction of the concept 'binge eating disorder' and the provision of research criteria for it. This has encouraged research on one subgroup of patients within eating disorder NOS. However, binge eating disorder comprises a small minority of cases within eating disorder NOS as indicated by the finding that the proportion of binge eating disorder cases in three well-diagnosed outpatient samples was in all three cases below 10% (Martin et al., 2000; Nollett & Button, 2005; Ricca et al., 2001; see also Chapter Three).

1.3.5.4 Criticisms of the DSM-IV scheme for classifying eating disorders

One major criticism of the DSM-IV scheme for classifying eating disorders stems from the aforementioned high relative prevalence of eating disorder NOS: the two specified eating disorder diagnoses, anorexia nervosa and bulimia nervosa, simply do not cover the eating disorder ground (Palmer, 2003). For the eating disorder ground to be covered, the diagnostic scheme for eating disorders has to rely on the residual category eating disorder NOS, a catch-all diagnosis that comprises approximately half of all eating disorder cases and as a diagnostic category does not inform treatment planning. Only having anorexia nervosa

and bulimia nervosa defined seems inadequate and a poor representation of clinical reality (Fairburn & Bohn, 2005).

Another weakness of the current scheme is that there is no definition of what is a "case". The boundaries around the two specified diagnoses anorexia nervosa and bulimia nervosa are defined, but that around the largest eating disorder category, eating disorder NOS (see Figure 1.1), is not. As pointed out earlier, this boundary is the boundary of eating disorder "caseness". In other words, a major problem with the current scheme for classifying eating disorders is the absence of an operational definition of what is an eating disorder. In Chapter Five a study is described that was designed to generate an empirically determined definition of what is an eating disorder.

A more fundamental criticism of the scheme concerns the whole basis for subclassifying the category "eating disorder" (Fairburn & Harrison, 2003; Waller, 1993; Walsh & Garner, 1997). The present scheme for classifying eating disorders into anorexia nervosa and bulimia nervosa is largely a historical accident (Fairburn and Bohn, 2005). It is not an empirically-derived scheme. Two observations raise doubts about the appropriateness of subdividing the eating disorders. First, and as discussed above, the three categories of eating disorder have many clinical features in common, most of which are not seen in other psychiatric disorders (e.g., strict and sustained attempts to restrict food intake; intermittent bouts of uncontrolled overeating (binge eating); self-induced vomiting; laxative misuse; driven over-exercising; judging self-worth almost exclusively in terms of shape and weight; pathological body checking and/or avoidance). And second, studies of their course have confirmed the clinical observation that many patients move between these diagnoses over time: indeed, temporal migration appears to be the norm rather than the exception. Unfortunately this diagnostic flux has received limited research attention (Sullivan, Bulik,

Carter, Gendall, & Joyce, 1996; Bulik, Sullivan, Fear, & Pickering, 1997; Keel, Fichter, Quadflieg, Bulik, Baxter, Thornton et al., 2004) despite being a well known phenomenon to clinicians. A recent exception is a study by Milos and colleagues (Milos, Spindler, Schnyder, & Fairburn, 2005). The authors studied the diagnostic flux in 192 individuals with an eating disorder of clinical severity. The sample was a mixture of inpatients, outpatients and eating disorder cases in the community. Participants were assessed at three time points (baseline, 12 months later and 30 month later) and eating disorder diagnoses were made at each time point. At baseline the diagnostic distribution was as follows: 55 (28.7%) fulfilled diagnostic criteria for anorexia nervosa, 108 (56.3%) had bulimia nervosa and 29 (15.1%) were diagnosed with eating disorder NOS. Twenty percent of the baseline anorexia nervosa cases moved over to bulimia nervosa at either or both follow-up points, and 9% migrated the other way. Over a third (37%) of all baseline anorexia nervosa and bulimia nervosa cases (N=163) were given a diagnosis of eating disorder NOS at either or both follow-up points. Twelve of the original eating disorder NOS cases (41%) moved over to a specified diagnosis (anorexia nervosa or bulimia nervosa) at either or both follow-up points. It was found that while the overarching diagnosis "eating disorder" was relatively stable across the three time points, there was considerable migration of subjects between the eating disorder diagnoses. A change in diagnosis occurred in over half the cases (53%), excluding those who no longer had an eating disorder at one or both follow-up points, and diagnostic migration was evident in all three diagnostic groups.

Recently these observations led Fairburn and colleagues to adopt a "transdiagnostic" perspective on eating disorders (Fairburn & Harrison, 2003) and their treatment (Fairburn, Cooper, & Shafran, 2003). It also let the candidate and Fairburn to propose a "transdiagnostic" solution to the nosological problem posed by eating disorder NOS

(Fairburn & Bohn, 2005). This was to create a single unitary diagnostic category "eating disorder" that would embrace anorexia nervosa, bulimia nervosa and eating disorder NOS without any subdivisions. The main argument for proposing a "transdiagnostic" solution of this type is that the current emphasis on subdividing the eating disorders (into anorexia nervosa and bulimia nervosa, each with their two subtypes, eating disorder NOS and possibly binge eating disorder) detracts attention from the most striking characteristic of the eating disorders; namely, that far more unites the various forms of eating disorder than separates them (Fairburn & Harrison, 2003; Waller, 1993; Walsh & Garner, 1997). Thus, rather than focusing on differences between the eating disorders, Fairburn and Bohn (2005) argue that there is a case for highlighting the many features that are shared by them and are largely peculiar to them. This, they believe, would encourage and permit the classification of eating disorders to be examined afresh. This is the premise that underpins this dissertation.

A small number of studies have attempted to derive clinically valid, natural subgroups within the overall category "eating disorder". All of them are, however, limited by the absence of an agreed definition on what constitutes an eating disorder of clinical severity. Sloan and colleagues (Sloan, Mizes, & Epstein, 2005) examined the DSM-IV classification of eating disorders using a cluster analytic approach with a sample of treatment-seeking patients diagnosed with anorexia nervosa, bulimia nervosa or eating disorder NOS. The variables used were the eight subscales of the Eating Disorder Inventory (EDI; Garner, Olmsted and Polivy, 1983) and a number of weight- and eating disorder behaviour-related variables (i.e., weight dissatisfaction, weight goal, weight fluctuation, current BMI, highest and lowest adult BMI, and the average number of self-reported binges and purges per week). Four cluster groupings were found. Cluster One appeared similar to

anorexia nervosa, both the binge/purge and restricting subtypes. Cluster Three resembled bulimia nervosa, as patients in this group were distinguished by a high frequency of bingeing and purging. Cluster Four resembled binge eating disorder in that there was a high frequency of binge eating and a high body weight. However, this subgroup also showed a high frequency of purging, which does not fit into the clinical picture of binge eating disorder. Cluster Two also seemed to resemble bulimia nervosa, but the patients had a lower, yet clinically significant, frequency of binge eating and purging than the patients in Cluster Three. The cluster groupings did not map on well to the DSM-IV diagnoses of anorexia nervosa, bulimia nervosa and binge eating disorder leading the authors to conclude that the specific DSM-IV criteria for eating disorders may need to be reconsidered. They suggested that current weight, weight history and frequency of bingeing and purging may be important in distinguishing subgroups of eating disorder patients since these variables appeared important in distinguishing the clusters.

Clinton and colleagues (Clinton, Button, Norring, & Palmer, 2004) also tried to find natural groupings of eating disorder patients. They used cluster analysis with two samples of adult female patients, one from Sweden and the other from England, and 10 key diagnostic clinical variables. These were selected from two eating disorder instruments, the Clinical Eating Disorders Rating Instrument (CEDRI; Palmer, Christie, Cordle, Davies, & Kenrick, 1987) and the Rating of Anorexia and Bulimia Interview (RAB; Clinton & Norring, 1999). The variables were BMI, fear of weight gain, restriction of food intake, avoidance of fattening foods, binge eating, self-induced vomiting, abuse of laxatives, compulsive exercise, amenorrhoea, and body-image disturbance. The authors found that a three-cluster solution provided the most parsimonious classification of the cases in both samples. The largest cluster was termed "generalised eating disorder" and was characterised

by high levels of eating disorder psychopathology on all key diagnostic variables except weight and menstrual functioning. The second cluster broadly resembled patients with DSM-IV anorexia nervosa (restricting type) in that it was characterised by low weight, amenorrhoea and the absence of binge eating. The third cluster was characterised by high weight and moderate levels of binge eating and compensatory behaviours. The distinction between the first cluster and the third seemed to be mainly based on severity and weight. The authors stated that, although the findings did not correspond precisely to the DSM-IV diagnoses, they broadly supported the distinction between anorexia nervosa (restricting type) and other eating disorders. In a subsequent study Clinton and Norring (2005) compared the clinical utility of these three statistically derived clusters to that of the three DSM-IV diagnoses using the Swedish sample of patients (see above). Comparisons were made with regard to eating disorder symptomatology and psychiatric comorbidity assessed at intake and subsequently after six and 36 months. The authors found that, compared to DSM-IV diagnoses, the three clusters demonstrated a higher degree of utility in terms of more distinct between-group differences and higher effect sizes in relation to a wide range of eating disorder psychopathology and psychiatric comorbidity. This was true at intake and at both follow-up points. The authors suggested that the ability of the cluster analytic approach to generate a more homogeneous and powerful classification may have been partly due its allocation of patients with eating disorder NOS to more relevant categories. The authors concluded that the predominance of the "generalised eating disorder" cluster "may in fact be an apt reminder that the time is ripe to seriously contemplate a move away from increased subcategorising of eating disorders toward defining a diagnosis of eating disorder per se" (Clinton & Norring, 2005, p.416). One serious limitation of this study is the high dropout rate (46% at 36-month follow-up) which means that its predictive findings need to be

interpreted with caution.

Several other studies have also attempted to derive more clinically meaningful eating disorder subgroups (Bulik, Sullivan, Kendler, 2000; Keel et al., 2004). However, these studies have relied upon life-time data which are of limited relevance to the classification of current symptomatology.

1.4 Eating Disorder Not Otherwise Specified (NOS)

1.4.1 Introduction

In this section of the chapter the DSM-IV diagnosis eating disorder NOS is reviewed with regard to its clinical characteristics, prevalence, course and treatment. As will become clear, the literature on eating disorder NOS is small, reflecting the neglect of the diagnosis and of patients who receive it. The literature on eating disorder NOS is also confusing, in that inconsistent labels and definitions of the diagnosis have been used. Eating disorder NOS has been termed "partial syndrome" (e.g., Martin et al., 2000; Woodside, Garfinkel, Lin, Goering, Kaplan et al., 2001), "atypical" (e.g., Cotrufo et al., 2004; Mitchell, Pyle, Hatsukami, & Eckert, 1986), "subthreshold" (e.g., Crow et al., 2002) and even "subclinical" (e.g., Cotrufo, Barretta, Monteleone, & Maj, 1998). These anorexia nervosa- and bulimia nervosa-centered definitions illustrate the focus that clinicians and researchers place on the two specified eating disorders. The term "subclinical" is a particularly inappropriate description of eating disorder NOS, given that patients with this diagnosis, by definition, have an eating disorder of clinical severity. The usefulness of the term "atypical" can be questioned too, given the high prevalence of eating disorder NOS (see Section 1.4.3 below).

1.4.2 Clinical features of Eating Disorder NOS

Not much is known about the clinical characteristics of eating disorder NOS. Case series and case reports indicate that they vary in form (Norvell & Cooley, 1986; Shisslak, Crago & Yates, 1989; Steiger & Ghadirian, 1989). Some cases of eating disorder NOS resemble anorexia nervosa or bulimia nervosa but do not quite meet their diagnostic criteria, either because one of the essential diagnostic features is missing (the term "partial syndrome" or "atypical" is sometimes used for these clinical states) or because one or more features fall below the specified threshold (here the term "subthreshold" seems appropriate) (Fairburn and Bohn, 2005).

Few studies have formally described the psychopathology of patients with eating disorder NOS. Most that exist have serious limitations for one or more of the following reasons: 1) they are studies of community samples with no formal assessment of impairment (something that is needed to make the diagnosis eating disorder NOS since it has to be established that the state detected is of clinical severity) (e.g., Woodside et al., 2001; Cotrufo et al., 1998); 2) they do not study the whole category eating disorder NOS, but only a subgroup within it (e.g., subthreshold anorexia nervosa or subthreshold bulimia nervosa) (e.g., Cotrufo et al., 2004; Crow et al., 2002; Herzog et al., 1993); or 3) they use weak measures, such as unvalidated interviews or self-report questionnaires (e.g., Martin et al., 2000; Nollett & Button, 2005). There are two exceptions, and they are described in detail below.

Ricca and colleagues (2001) assessed 189 female patients with an eating disorder of clinical severity who were seeking treatment at one of two outpatient eating disorder clinics in Florence. The diagnostic distribution of the sample was as follows: 47 patients (24.9%) fulfilled DSM-III-R diagnostic criteria for anorexia nervosa, 47 (24.9%) were diagnosed

with DSM-III-R bulimia nervosa, and 95 (50.3%) were diagnosed with eating disorder NOS. Sixteen of the eating disorder NOS cases (8.5% of the total sample) fulfilled DSM-IV research criteria for binge eating disorder. The three groups were compared with respect to their demographic features, their eating disorder psychopathology (as assessed using a validated Italian translation of the EDE (Mannucci, Ricca, Di Bernardo & Rotella, 1996)) and their general psychopathology (as assessed using the Structured Diagnostic Interview for DSM-III-R (SCID; Spitzer, Williams, Gibbon & First, 1992), the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock & Erbaugh, 1961) and the State and Trait Anxiety Inventory (STAI; Spielberger, 1983). The only statistically significant differences that emerged between the three diagnostic groups were the age of onset of the eating disorder and the body mass index. Patients with eating disorder NOS were significantly older than those with anorexia nervosa at the time of onset of the disorder, but not in comparison with those with bulimia nervosa. Patients with eating disorder NOS had a higher body mass index than those with anorexia nervosa, a difference to be expected given the definition of the two disorders. There were no significant differences between the three disorders with regard to the lifetime prevalence of other psychiatric disorders, their current eating disorder psychopathology (global EDE score) and their current general psychopathology (total scores on STAI and BDI). On the basis of these findings, Ricca et al. (2001) suggested that the three eating disorders all have similar specific and general psychopathological characteristics. Their main conclusion was as follows: "The absence of any significant differences in the psychopathological features of EDNOS patients and those with formal eating disorders (...) suggests that they may experience a similar level of distress, which is why we feel that EDNOS should not be considered a "minor" or less severe form of eating disorder and that its clinical relevance should not be underestimated." (Ricca et al., 2001).

Using the EDE interview (Fairburn & Cooper, 1993), Turner and Bryant-Waugh (2004) assessed the eating disorder psychopathology of 190 adult patients referred to an outpatient-based eating disorder service. They compared the EDE profiles of three diagnostic groups, anorexia nervosa, bulimia nervosa and eating disorder NOS. Eleven (5.8%) fulfilled diagnostic criteria for anorexia nervosa, 45 (23.7%) were diagnosed with bulimia nervosa and 134 (70.5%) received the diagnosis eating disorder NOS. The authors found no statistically significant differences between the three groups in relation to the Dietary Restraint subscale of the EDE. The eating disorder NOS patients did not differ from the anorexia nervosa patients on any of the other three subscales (Eating Concern, Weight Concern, and Shape Concern): indeed, they only differed from the bulimia nervosa patients on a few individual EDE items. The three groups also did not differ with regard to age. The anorexia nervosa patients had a lower BMI than the other two groups, and the patients with eating disorder NOS were lighter than the bulimia nervosa patients. Turner & Bryant-Waugh (2004) concluded that patients with the diagnosis eating disorder NOS are comparable to anorexia nervosa and bulimia nervosa patients with regard to the severity of their eating disorder psychopathology, and that they merit more attention in research and clinical practice.

The findings from these two studies indicate that eating disorder NOS comprises a large group of patients in adult outpatient settings. Furthermore, it seems that the eating disorder features and general psychopathology of these patients are comparable in character and severity to those of patients with anorexia nervosa and bulimia nervosa. However, the two studies have a number of limitations. Ricca and colleagues used a sample that was not catchment-area based but consisted of patients who were seeking treatment in one of two clinics in Florence. Its representativeness could therefore be questioned. Although Turner

and Bryant-Waugh (2004) described the eating disorder psychopathology of their patients in great detail, they did not report on their general psychopathology or their histories. Also they excluded cases with comorbid obesity and thus many of those with binge eating disorder. In conclusion, further research on the characteristics of patients with eating disorder NOS is needed using representative patient samples and standardised measures of a broad range of psychopathology.

Several studies have attempted to identify subgroups within eating disorder NOS. Using a cluster analytic approach Mizes and Sloan (1998) found one main subgroup characterised by a high body mass index, a history of large weight fluctuations, and a high frequency of binge eating. Although these individuals seemed to resemble those with binge eating disorder, they also reported inappropriate compensatory behaviour and thus differed from those with binge eating disorder. Mizes and Sloan found another subgroup which was heterogeneous in character. A limitation of their study was that the sample consisted of a mixture of inpatients and out-patients.

Williamson and colleagues also used a cluster analytic approach to look for subgroups within eating disorder NOS (Williamson, Gleaves & Savin, 1992). They found three distinct subgroups that they described as subthreshold anorexia nervosa, bulimia nervosa (non-purging type) and a group that resembled binge eating disorder. This study had three serious limitations: first, similar to the Mizes and Sloan study, the eating disorder NOS patients were a mixture of inpatients and out-patients. Second, the definition of eating disorder NOS was problematic in that patients were only included if they "reported most but not all of the symptoms associated with anorexia nervosa or bulimia nervosa". Thus this sample consisted of a subgroup of patients within eating disorder NOS whose clinical state was of uncertain clinical significance as there was no assessment of secondary impairment.

Third, the authors excluded cases of non-purging bulimia nervosa from the diagnosis of bulimia nervosa, therefore artificially inflating the eating disorder NOS category.

Keel and colleagues recently proposed demarcating a new category of eating disorder from within the eating disorder NOS group (Keel, Haedt & Edler, 2005). This is characterised by recurrent purging in the absence of regular binge eating. They termed the state "purging disorder". Keel and colleagues argued that purging disorder should be recognised as a distinct eating disorder as they found that patients meeting these criteria showed greater eating and general psychopathology than controls and lower eating concerns, disinhibition and hunger than cases of bulimia nervosa. This said, overall these cases were very similar to those with bulimia nervosa; for example, their eating and general psychopathology was indistinguishable from the cases of bulimia nervosa. Originally Keel and colleagues named this state "subjective bulimia nervosa" (Keel, Mayer & Hamden-Fischer, 2001) and given its similarity to bulimia nervosa this might have been a better term. Certainly at present there is not enough evidence to support recognising purging disorder as a distinct eating disorder.

1.4.3 Prevalence of Eating Disorder NOS

As pointed out in section 1.3.5.3 of this chapter, eating disorder NOS is the most common diagnosis made in most adult outpatient settings other than those that attract highly specialist referrals. Table 1.6 in Section 1.3.5.3 of this chapter shows the prevalence figures from five well-diagnosed samples of adults seeking treatment for an eating disorder. In each case eating disorder NOS was the most common diagnosis made, its weighted average being 54.2%.

Studies of the prevalence of eating disorder NOS in the community have one limitation in common; the lack of any assessment of secondary impairment, something that

is essential for the diagnosis of eating disorder NOS to be made since, as noted above, it has to be established that the state detected is of clinical severity. The prevalence figures quoted tend to be for the prevalence of participants with features suggestive of an eating disorder (other than anorexia nervosa or bulimia nervosa), no check being made that these features result in a clinically significant level of secondary impairment (e.g., Cotrufo et al., 1998; Garfinkel et al., 1995b; Hay, Fairburn & Doll, 1996; Woodside et al., 2001). The findings of these studies are therefore impossible to interpret.

1.4.4 Course of Eating Disorder NOS

While many studies have investigated the course and outcome of patients with anorexia nervosa and bulimia nervosa, very few have focused on eating disorder NOS. An exception was the study of Herzog and colleagues (1993). They studied 33 cases of "subdiagnostic anorexia nervosa" and "subdiagnostic bulimia nervosa" over a period of 24 to 52 months after they first sought treatment for their eating disorder. They found that these cases showed a "persistent and highly variable eating disorder course". Forty-six percent went on to meet full DSM-III-R criteria for anorexia nervosa or bulimia nervosa during the period of follow-up. The recovery rate was low at 18%. This study is the first to examine the clinical course of patients with eating disorder NOS. However, due to its exclusive focus on subthreshold forms of anorexia nervosa and bulimia nervosa, it only studied a small proportion of those with eating disorder NOS (see study described in Chapter Four).

Strober and colleagues (1999) studied the natural course of 77 patients with DSM-III anorexia nervosa and 20 patients with "atypical anorexia nervosa" over a time period of 10 to 15 years after inpatient treatment. The atypical cases fulfilled all diagnostic criteria for anorexia nervosa with the exception of fear of weight gain and body image distortion. The authors found that, in comparison with the anorexia nervosa cases, the atypical cases had a

faster rate of full clinical recovery and a lower risk of developing binge eating. It is important to note that this study only examined the course of a subgroup of eating disorder NOS cases.

A recent study that investigated the course of the three eating disorders was described earlier (see Section 1.3.5.4). Milos and colleagues (2005) examined the diagnostic flux in 192 individuals with an eating disorder of clinical severity, 29 of which (15.1%) were diagnosed with eating disorder NOS. The patients were assessed at three time points (baseline, 12 months later and 30 month later) and eating disorder diagnoses were made at each point. At the 12-month assessment 69% of the original eating disorder NOS cases had either developed anorexia nervosa or bulimia nervosa or had retained their diagnosis, whereas 31% no longer had an eating disorder. At the 30-month assessment 48% of the original eating disorder NOS still had an eating disorder, whereas 52% did not.

The findings of this study have to be interpreted with caution, given the following three important limitations: First, the eating disorder NOS sample was very small in size, second the participants were a mixture of inpatients, outpatients and eating disorder cases in the community, and third, most of the patients had received treatment between assessments which might have influenced the course of the eating disorder. It is therefore difficult to generalise from this small and possibly unrepresentative sample to eating disorder NOS as a whole.

1.4.5 Treatment of Eating Disorder NOS

The research on the treatment of eating disorders has focused on anorexia nervosa and bulimia nervosa. There are studies of the treatment of subgroups of eating disorder NOS, e.g., of patients with subthreshold bulimia nervosa (Nevonen & Broberg, 2005) or binge eating disorder (e.g., Marcus, Wing & Fairburn, 1995; Wilfley & Cohen, 1997;

Wilfley, Welch, Stein, Spurrell, Cohen et al., 2002). As will be clear from the findings presented in Chapter Four, such patients are a small proportion of those with eating disorder NOS. To date no research on the treatment of the whole group of eating disorder NOS has been conducted (although one in Oxford is nearing completion). This highlights the neglect of this diagnostic category and of patients who receive the diagnosis.

The lack of research on the treatment of patients with eating disorder NOS has far-reaching implications for research and clinical practice. First, it leads to the potential inflation of recovery rates of anorexia nervosa and bulimia nervosa as eating disorder NOS has been excluded as an outcome in treatment studies. Secondly, it leaves clinicians in the dark about how to best treat the most common group of eating disorder patients. The UK National Institute for Clinical Excellence (NICE) recently published clinical guidelines for the treatment of eating disorders (National Collaborating Centre for Mental Health, 2004). The guidance is evidence-based and the recommendations are graded into three levels: an "A" recommendation means that it is supported by good evidence (i.e., at least one randomised controlled trial (RCT)), a "B" recommendation means there is some supporting evidence (i.e., good clinical studies, but no RCT), and a "C" recommendation means there is no empirical evidence other than expert opinion (i.e., no directly applicable clinical studies of good quality). The recommendation for the treatment of patients with eating disorder NOS (a "C") provided by NICE was as follows:

"In the absence of evidence to guide the management of atypical eating disorders (eating disorders not otherwise specified) other than binge eating disorder, it is recommended that the clinician considers following the guidance on the treatment of the eating problem that most closely resembles the individual patient's eating disorder."

(National Collaborating Centre for Mental Health, 2004, p.71)

Whilst logical, this recommendation is vague and supported by no evidence.

Clearly the current situation is unsatisfactory. The eating disorder field needs research on the most common eating disorder diagnosis, eating disorder NOS, especially with regard to its treatment. Research on eating disorder NOS is, however profoundly limited by the absence of a consistent definition of the diagnosis. To derive one requires the identification of the type and level of eating disorder features that are associated with clinically significant impairment. This topic is addressed in the next section of this chapter.

1.5 Clinical significance

1.5.1 Introduction

As has been pointed out earlier in the chapter, DSM-IV (APA, 1994) requires that there be clinically significant distress or impairment in social, occupational, or other areas of functioning secondary to the specific behavioural or psychological disturbance, for a diagnosis of a mental disorder to be given. For that reason a clinical significance criterion was added to the criteria sets of most disorders, helping to establish the threshold between disorder and abnormal (but subthreshold) behaviour or concerns. This was done since there was concern about clinicians applying diagnostic criteria in a 'cookbook fashion', without checking whether the identified features were actually causing significant impairment or distress (APA, 1994). The inclusion of clinical significance criteria was a positive step towards a better delineation between pathological states and more normative psychological, behavioural and emotional features. However, determining what constitutes 'clinically significant distress or impairment' remains an inherently difficult clinical judgment.

DSM-IV includes a scale for reporting a clinician's judgment of an individual's overall level of functioning, the "Global Assessment of Functioning" (GAF) scale on Axis V (APA, 1994, p.32). The GAF scale ranges from 0 to 100 with higher scores indicating better

functioning. It is designed to be rated with respect to the individual's psychological, social, and occupational functioning although no guidance is provided as to how the individual's functioning should be assessed. The GAF's potential as a measure of clinical impairment is limited for two reasons. First, the rating of functioning is confounded with the rating of psychiatric symptoms (i.e., contributing to the GAF score are symptoms) with the result that the scale fails to measure uncontaminated functioning. Secondly, the GAF does not ensure that any impairment of functioning detected is secondary to the psychiatric features in question. As Kendler pointed out: "The multiaxal DSM-IV system captures functioning on axis V, but there is no attempt to determine whether dysfunction is 'due to' the diagnosed disorder or disorders" (Kendler, 1999, p.1847).

The eating disorder field needs a means for identifying clinically significant impairment due to the presence of eating disorder psychopathology. This is because, as pointed out earlier, the most common eating disorder diagnosis made in outpatient settings is eating disorder NOS, a diagnosis with no other diagnostic criteria than the requirement that there be an eating disorder of clinical severity (i.e., the eating disorder features present need to result in clinically significant impairment) and that the diagnostic criteria for anorexia nervosa and bulimia nervosa are not met. Therefore a measure of impairment secondary to eating disorder features is needed. The next section of this chapter reviews existing measures of impairment and health-related quality of life in people with eating disorders.

1.5.2 Impairment in the eating disorders

Despite the need to describe and measure the impairment that is secondary to eating disorder psychopathology, there have been few attempts to do so. Instead of impairment per se, researchers and clinicians have begun to study the 'quality of life' and the 'health related quality of life' of individuals with eating disorders. Quality of life (QoL) refers to an

individual's overall satisfaction with his or her life (Kushner & Foster, 2000). It encompasses a broad range of dimensions, including the individual's subjective physical, emotional, social, sexual, and occupational functioning (Spilker & Revicki, 1996). Health-related quality of life (HRQoL) does not have a specific definition as such although there is agreement that it is a multidimensional construct and concerns the impact of a specific illness on three major aspects of functioning: physical, psychological, and social (Boini, Briancon, Guillemin, Galan & Hercberg, 2004). Instruments measuring HRQoL assess the impact of specific illnesses and their treatment on the subjective evaluation of various health domains. Both QoL and HRQoL are concepts that have become important outcome variables in many areas of clinical research. In the eating disorder field, however, they have been largely ignored. Padierna and colleagues pointed out in 2000 that there is a distinct lack of relevant literature on health-related quality of life research specific to eating disorders despite the disorders' being associated with significant physical, psychological and social dysfunction (Padierna, Quintana, Arostegui, Gonzalez & Horcajo, 2000).

Since then a small number of studies have examined (health-related) quality of life in individuals with eating disorders. Most of these studies have used generic (health-related) quality of life measures, such as the 'Medical Outcomes Study 36-item Short-Form Health Survey' (SF-36; Ware, Snow, Kosinski & Reese, 1993) or the 'World Health Organisation Brief Quality of Life Assessment Scale' (WHOQOL-BREF; WHOQOL Group, 1998).

The SF-36 is a 36-item health questionnaire designed to provide a comprehensive measure of physical, emotional, and social well-being. It focuses on functioning in a number of life domains over the past four weeks. It includes eight subscales: General Health, Physical Functioning, Role-Physical (physical limitations), Bodily Pain, Mental Health, Role-Emotional (limitations from emotional difficulties), Vitality, and Social Functioning.

SF-36 scores range from 0 to 100 with a higher score indicating better quality of life. A 12-item version has been derived from the SF-36: the SF-12 (Ware, Kosinski & Keller, 1996). Items of the SF-12 are summarised into two weighted scales (the Physical Component Summary scale and the Mental Component Summary scale), designed to assess impairment in everyday functioning associated with physical and mental health problems.

The World Health Organisation Brief Quality of Life Assessment Scale (WHOQOL-BREF; WHOQOL Group, 1998) is a 16-item short-form of the WHOQOL-100 assessment scale (WHOQOL Group, 1998). It yields scores on four life domains related to the individual's subjective evaluation of quality of life: physical health, environmental health, psychological health, and social relationships. Items (and subscales) are scored on a scale from 1 to 5, with a score of "1" indicating extreme dissatisfaction and a score of "5" indicating extreme satisfaction.

The few studies in the eating disorder field that have used the SF-36, the SF-12, or the WHOQOL-BREF have found that the health-related quality of life is significantly impaired in patients with eating disorders (Padierna et al, 2000; Masheb & Grilo, 2004; Mond, Hay, Rodgers, Owen & Beumont, 2005) and in eating disorder cases in the community (Hay, 2003; Mond, Rodgers, Hay, Korten, Owen & Beumont, 2004; Doll, Petersen & Stewart-Brown, 2005; De la Rie, Noordenbos & van Furth, 2005) in comparison with that of people without eating disorders. The impairment affects all aspects of life. It has also been found that the level of eating disorder features is correlated with the level of impairment (Padierna et al., 2000). All the studies, however, have had one or more of the following limitations. First, very few studied individuals with all three eating disorder diagnoses (Mond et al., 2004; De la Rie et al., 2005). Most have either focused on cases of anorexia nervosa, bulimia nervosa and binge eating disorder (Padierna et al., 2000; Mond et

al., 2005; Doll et al., 2005) or only on binge eating disorder (Masheb & Grilo, 2004), therefore not studying the whole group of individuals with eating disorders (i.e., they have omitted most cases of eating disorder NOS). Second, all the studies have used generic HRQoL measures. Such measures allow comparisons of quality of life across various disorders but they may not be sensitive to the domains of life that might be specifically affected by eating disorders. Third, and possibly most importantly, the measures used have failed to ensure that the impairment is secondary to the person's specific disorder; in this case, their eating disorder. It cannot be said, for example, that the impairment is not due to the frequently occurring, comorbid psychiatric disorders (e.g., clinical depression) or to negative life events. This limitation makes the existing generic health-related quality of life measures unsuitable for the assessment of impairment secondary to eating disorder features.

A development in the obesity field was the creation of a weight-related quality of life measure, the Impact of Weight on Quality of Life (IWQOL; Kolotkin, Head, Hamilton & Tse, 1995), which was subsequently used with patients with binge eating disorder (Kolotkin, Westman, Ostbye, Crosby, Eisenson & Binks, 2004; Rieger, Wilfley, Stein, Marino & Crow, 2005). This measure addresses the second and third limitations identified above. The short version of the IWQOL, the IWQOL-Lite (Kolotkin, Crosby, Kosloski & Williams, 2001), has been used in the studies of binge eating disorder. It is a 31-item, self-report questionnaire that assesses the effect of being overweight on five domains of life: Work, Public Distress, Sexual Life, Physical Function, and Self-Esteem. Scores on each subscale are summed to yield a total score, with higher scores indicating greater impairment. This instrument differs from generic health-related quality of life instruments in two ways: first, it assesses impairment in areas of life that are relevant to and commonly affected by the

specific condition (in this case, obesity), and second, it is ensured that the impairment is attributable to the specific condition (and not to comorbid conditions or life events).

A very recent development in the eating disorder field has been the creation of eating disorder specific health-related quality of life instruments. They are designed to focus on the domains, characteristics and complaints most relevant to people with eating disorders. Three such measures have been developed and they will be described in turn.

Las Hayas and colleagues describe the development and psychometric properties of the "Health-Related Quality of Life in Eating Disorders" (HeRQoLED) questionnaire (Las Hayas, Quintana, Padierna, Bilbao, Munoz et al., 2006). The instrument consists of 50 items. It has eight health-related domains that were chosen on the basis of discussions amongst eating disorder sufferers and their family members and eating disorder professionals regarding which aspects of life are most commonly affected by eating disorders. The eight domains are: Physical Symptoms, Physical Role, Restrictive Behaviours, Body Image, Mental Health, Emotional Role, Social Relations, and Personality Traits. Each item is rated on a 0 to 4 or 5 scale, with higher scores indicating lower perception of quality of life. The total score for each domain-related subscale is computed by adding the scores of each item assigned to the domain and standardising the result on a scale of 0 to 100. The instrument does not generate a total score although one could be computed. Unfortunately, the instrument's potential as a measure of secondary impairment is limited. First, with the exception of the Social Relations domain, no check is made whether the impairment is due to the person's eating disorder features. Second, the measurement of health-related quality of life is confounded with the measurement of eating disorder symptoms (i.e., contributing to the impairment scores are eating disorder features)

with the result that the ability of the instrument to measure uncontaminated secondary impairment is profoundly limited.

Abraham and colleagues (Abraham, Brown, Boyd, Luscombe & Russell, 2006) describe the development of the "Quality of Life for Eating Disorders" measure (QOL ED), a computer-administered, -scored and -reported instrument derived from the Eating and Exercise Examination (EEE-C; Abraham & Lovell, 1999). The QOL ED consists of 21 questions. Six scores can be derived from its questions: Body Weight, Eating Behaviour, Eating Disorder, Psychological, Daily Living (perceived effect on study/work, social life and relationships), and Acute Medical. There is also a global QOL ED score, which is the total of all the six above scores and ranges from 0 to 24. Higher scores indicate a greater impairment of eating disorder-specific health-related quality of life. Two time frames exist; the past 28 days and the past three months. The instrument showed good reliability and sensitivity to change, and differentiated between cases of anorexia nervosa, bulimia nervosa and eating disorder NOS. In addition, the domain scores were significantly correlated with measures of eating disorder features, psychological dysfunction, general quality of life and body weight. With regard to the ability to assess impairment due to the person's eating disorder features, however, this instrument also has limitations. Like the instrument developed by Las Hayas and colleagues (2006), the QOL ED confounds eating disorder features and indices of impairment and thus some scores are a mixture of the two. It assesses a mixture of eating disorder symptoms (Eating Behaviour and Eating Disorder), general psychological symptoms that cannot be directly attributed to the person's eating disorder (Psychological), and impairment secondary to the person's "eating, exercise or body weight" (Daily Living and Acute Medical). The instrument's ability to detect impairment

due to eating disorder features is therefore limited to the person's social life and their physical health.

The third eating disorder specific HRQoL instrument is the "Eating Disorders Quality of Life" (EDQOL; Engel, Wittrock, Crosby, Wonderlich, Mitchell & Kolotkin, 2006)). The EDQOL was developed by a team of six eating disorder experts. First, on the basis of their clinical experience, they identified areas of health-related quality of life most commonly affected by eating disorders. These areas were physical, psychological, financial, social, work/school, and legal. Then, content and items were generated for each domain. An exploratory factor analysis was then conducted on a sample of 538 adult women (155 with an eating disorder (mainly college students), 327 students without an eating disorder and 56 students who showed some eating disorder features. It revealed the presence of four factors which were used to create the 25-item instrument consisting of four subscales: Psychological (nine items), physical/cognitive (six items), financial (five items), work/school (five items). Each item is scored on a 0 to 4 scale, ranging from "Never" to "Always", and subscale scores are generated by averaging the items of each scale. A total score is calculated by averaging the scores of all of the items and also ranges from 0 to 4 with lower scores indicating a better health-related quality of life. The instrument focuses on the person's state over the previous 30 days. The authors examined the convergent validity of the EDQOL by comparing the performance of each subscale with an existing self-report measure of the same construct or a newly developed one ("Financial Global Ratings"). The domains showed high correlations with their respective measures (BDI [Beck, 1961]; SF-36 [Ware et al., 1993]; "Neuroticism" as measured by Goldberg, 1992; Work subscale of the SAS-SR [Weissman & Bothwell, 1976]). It was found that the participants with an eating disorder scored significantly higher on all four subscales and had a higher total score than

the other two groups, and that more severe eating disorder psychopathology was associated with higher EDQOL scores. With regard to its reliability, the instrument appeared to be internally consistent and demonstrated good test-retest reliability, with the exception of the Work/School subscale.

Unlike the other measures, the EDQOL ensures that the functional impairment assessed is secondary to the person's eating disorder features. Each question starts with the phrase "How often has your eating/weight resulted in (or caused/affected/made you) ...", therefore requiring the participant to focus exclusively on impairment in that domain that is due to their "eating/weight". A serious limitation of the instrument is, however, that it will not detect the full extent of the secondary psychosocial impairment since its stem question neglects the one component of eating disorder psychopathology that is most troublesome to most patients, their overconcern about their shape. As a result the EDQOL is likely to markedly underestimate the extent of the secondary impairment.

1.5.3 Conclusions and implications

When assessing psychiatric disorders it is not sufficient to just measure symptoms. The impact of these symptoms on the person's life also needs to be considered. Many measures of eating disorder features exist but, as yet, there is no satisfactory measure fully capturing impairment of functioning secondary to them. In Chapter Six the development and validation of such a measure is described.

In addition, impairment needs to be assessed to determine eating disorder "caseness" as no symptomatic severity thresholds are specified in DSM-IV except for anorexia nervosa and bulimia nervosa (i.e., the two least common diagnoses) and these thresholds have been criticised for being arbitrary and inappropriate. In Chapter Five a transdiagnostic

impairment-based definition of eating disorder "caseness" is derived using an interview-based measure of impairment.

1.6 Summary

1. Three eating disorders are recognised in DSM-IV (APA, 1994). The diagnostic criteria for the two specified eating disorders, anorexia nervosa and bulimia nervosa, have been criticised for being arbitrary, too strict and not based on empirical evidence. The third eating disorder diagnosis, "eating disorder NOS", a residual category, is the most common eating disorder diagnosis made in most adult outpatient settings. There is clearly something amiss with a diagnostic system in which the most common category is the supposedly residual one. Possible solutions to these inter-related problems are presented in Chapter Four.
2. The diagnosis eating disorder NOS has been neglected by researchers and clinicians alike. Even its clinical features have been poorly defined. Chapter Three describes a study of its clinical features.
3. Positive diagnostic criteria are needed for the diagnosis eating disorder NOS. To devise such criteria an operational definition is needed of what is an "eating disorder". This in turn requires the identification of type and level of eating disorder features that are associated with clinically significant impairment of functioning. A study that addresses these issues is described in Chapter Five.
4. For clinical and research purposes it would be of great value to have a readily administered measure of psychosocial impairment secondary to eating disorder features. In Chapter Six the development of such a measure is described.

CHAPTER TWO

Aims of the Research and Outline of the Dissertation

2.1 General aims

The research described in this dissertation had four overarching aims. The first was to describe the clinical characteristics of people given the psychiatric diagnosis "Eating disorder not otherwise specified" (Eating disorder NOS), the residual, yet the most common of the three eating disorder diagnoses listed in the latest edition (fourth) of the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders (APA, 1994). The second was to examine how the problems of nosology and neglect associated with this diagnosis could potentially be solved. The third was to derive a provisional operational definition of what constitutes an eating disorder of clinical severity. The fourth aim was to develop a clinically useful self-report measure of psychosocial impairment secondary to eating disorder psychopathology.

2.2 Specific aims

The four general aims each encompass a number of more specific aims. These are listed below.

Aim One: To describe the clinical characteristics of people with eating disorder NOS

a. To compare the clinical features of a sample of adult outpatients with eating disorder NOS with those of outpatients with bulimia nervosa, thereby determining the character and severity of eating disorder NOS

b. To characterise in detail the most common behavioural eating disorder features of patients with eating disorder NOS and to determine the most common subgroups

Aim Two: To examine how the problems of nosology and neglect associated with eating disorder NOS could be solved

a. To propose new ways of conceptualising the diagnosis eating disorder NOS that would address its neglect and anomalous nosological status

b. To examine the impact of one of the proposed re-conceptualisations on the relative prevalence of eating disorder NOS.

c. To evaluate the clinical utility of another suggested re-conceptualisation of eating disorder NOS.

Aim Three: To derive a provisional operational definition of what constitutes an eating disorder of clinical severity

a. To develop an interview measure of functional impairment secondary to eating disorder psychopathology

b. To examine the psychometric properties of the measure

c. To derive a set of clinically valid positive diagnostic criteria for an 'eating disorder', utilising the newly developed measure of impairment, thereby identifying the boundary between eating disorder 'caseness' and lesser, not clinically significant, problems with eating

Aim Four: To develop a clinically useful self-report measure of psychosocial impairment secondary to eating disorder psychopathology

a. To examine the psychometric properties of the measure

2.3 Outline of the dissertation

Each general aim is addressed in a separate chapter. The content of each chapter is briefly outlined below.

Chapter Three addresses the first of the four general aims. It describes:

- A study which characterises the clinical features of patients with eating disorder NOS and compares them with those of patients with bulimia nervosa, using a large and representative sample of adult outpatients.
- A study which characterises in detail the behavioural eating disorder features of patients with eating disorder NOS and determines the most common subgroups

Chapter Four addresses the second general aim. It describes:

- Three proposals developed by the candidate and her supervisor (Fairburn and Bohn, 2005) for reconceptualising the diagnosis eating disorder NOS in order to address the problems of nosology and neglect associated with the diagnosis.
- A study which examines the impact of the first of Fairburn and Bohn's proposals on the relative prevalence of eating disorder NOS, using a representative sample of adult patients with all forms of eating disorder.
- A study which examines the 'clinical utility' (as defined by First et al (2004)) of the second of Fairburn and Bohn's proposals by surveying the opinions of eating disorder experts by means of a self-report questionnaire.

Chapter Five addresses the third general aim. It describes:

- The development of an interview-based measure designed to assess the clinical significance of any psychosocial and physical impairment secondary to eating disorder features.

- Three studies of the reliability (inter-rater and test-retest) and the validity (construct and discriminant) of this interview.
- A study of a large and representative sample of patients with all forms of eating disorder, assessed prospectively with the interview and a leading measure of eating disorder psychopathology, that was designed to identify the type and level of eating disorder features associated with a clinically significant level of secondary psychosocial or physical impairment.

Chapter Six addresses the fourth general aim. It describes:

- The development of a readily-used self-administered measure of psychosocial impairment secondary to eating disorder features
- Four studies of this instrument's properties; namely, its reliability, validity and sensitivity to change.

Chapter Seven

In this chapter the overall aims of the dissertation are restated and the findings are summarised. The conclusions that can be drawn from the research and the broader relevance of the work are discussed. Finally, future directions for research are proposed

CHAPTER THREE

The Clinical Features of Eating disorder NOS

3.1 Introduction

As discussed in Chapter One few studies have formally described the clinical features of patients with eating disorder NOS, and most that exist have serious limitations (see Chapter One, Section 1.4.2). Only two studies exist that have examined a well-characterised group of outpatients with eating disorder NOS using standardised eating disorder measures (Ricca et al., 2001; Turner & Bryant-Waugh, 2004). From these studies it is known that patients with eating disorder NOS resemble those with anorexia nervosa and bulimia nervosa with regard to their demographic features, their eating disorder psychopathology and their general psychopathology. However, the two studies showed a number of limitations. Ricca and colleagues used an unrepresentative sample, and Turner & Bryant-Waugh did not assess the general psychopathology of their representative sample of patients with eating disorders (see Chapter One, Section 1.4.2). Furthermore Turner & Bryant-Waugh excluded patients with binge eating disorder. Further research on the characteristics of patients with eating disorder NOS is therefore needed using representative patient samples and standardised measures of a broad range of psychopathology.

This chapter is concerned with examining and describing in detail the clinical characteristics of patients with the DSM-IV diagnosis eating disorder NOS using a large and representative clinical sample of outpatients with the diagnosis. There are two sections. In section 3.2 both the full range of eating disorder features and the general psychopathology of patients with eating disorder NOS are described and compared with those of patients with bulimia nervosa, an eating disorder of established severity. In section 3.3 the key behavioural features of the same sample of eating disorder NOS patients are examined

further. The most common behavioural subgroups within eating disorder NOS are identified.

3.2 The Clinical Features of Eating Disorder NOS in Comparison with Bulimia Nervosa

3.2.1 Aim

The present study had two aims. The first was to describe the clinical characteristics of eating disorder NOS by assessing a large and representative patient sample using standardised instruments. The second was to establish the severity of eating disorder NOS by comparing these cases with those with bulimia nervosa, the second most common eating disorder and one of established severity.

3.2.2 Methods

Design

To achieve these two aims the candidate was able to make use of a transdiagnostic clinical sample of outpatients with eating disorders that had been recruited in Oxford and Leicester for a treatment trial funded by the Wellcome Trust[2] For the purposes of the studies described in this chapter only the cases of eating disorder NOS and bulimia nervosa were relevant. However, for the purpose of studies later in this dissertation the entire sample is of relevance. Therefore, to avoid repetition, the full sample is described in this chapter.

The sample was recruited at two eating disorder outpatient clinics in the UK, one serving central Oxfordshire (the Centre for Research on Eating Disorders in Oxford (CREDO) based at Oxford University Department of Psychiatry) and the other serving Leicester and surrounding area (the Eating Disorder Unit at Leicester Hospital). Both clinics provide the only secondary eating disorder service for the locality. The patients included in

[2] The candidate was a research therapist at the Centre for Research in Eating Disorders in Oxford from October 2001 to September 2006.

this study were consecutive referrals to either of the two services, referred by their general practitioner or another clinician for the treatment of an eating disorder. Patients were assessed by a senior clinician (Dr Zafra Cooper or Professor Christopher Fairburn in Oxford; Professor Robert Palmer in Leicester) who determined whether the patient had an eating disorder of clinical severity and, if applicable, obtained a history of the eating disorder. Participants were subsequently assessed by a trained research assistant (Marianne O'Connor, Caroline Plumb and Shani Langdon in Oxford; Jackie Wales, Claire Nollett and Elizabeth Benson in Leicester) using the various measures listed below. DSM-IV eating disorder diagnoses were assigned using the Eating Disorder Examination (EDE) interview (Fairburn & Cooper, 1993) (see below).

Participants

The patient sample comprised consecutive referrals to the two eating disorder clinics described above. Patients were included in the transdiagnostic treatment trial if they met the following criteria: aged 18 to 65 years, judged to have a clinical eating disorder by one of three senior specialists in the field, body mass index between 16.0 and 39.9, able to attend for 20 or 40 sessions of outpatient treatment, and judged to be safe to manage on an outpatient basis. Each referral was evaluated by one of the three senior specialists who established whether the patient met the inclusion criteria listed.

Measures

Eating disorder features

Eating Disorder Examination 15.0 (EDE; Cooper & Fairburn, 1987; Fairburn & Cooper, 1993)

This interview was used to describe and measure the psychopathology that is specific to eating disorders and to generate DSM-IV diagnoses. The EDE is a semi-structured investigator-based interview and is currently regarded as the gold-standard for assessing

eating disorders (Garner, 1995). The instrument has gone through numerous revisions and editions, and has become the leading assessment instrument in studies of eating disorders (e.g., Agras, Walsh, Fairburn, Wilson & Kraemer, 2000; Fairburn, Cooper, Doll, Norman & O'Connor, 2000). The EDE can be used to generate specific eating disorder diagnoses and it assesses the behavioural and attitudinal psychopathology of eating disorders. With the exception of the diagnostic items, which are assessed and rated for a period of three months before assessment, the EDE focuses on the preceding 28 days. It assesses eating and meal patterns, the frequency of specific forms of overeating, including binge eating (objective bulimic episodes), and the use and frequency of different forms of weight-control behaviour (self-induced vomiting, laxative misuse, diuretic misuse, diet pill misuse, and fasting). Four subscales may be derived from the EDE: Dietary Restraint, Eating Concern, Weight Concern, and Shape Concern. Most items are rated on a 7-point scale (0-6), and raw 28-day frequencies are used for the behavioural items. The Global EDE score is the average of the four subscale scores and is an index of the overall severity of eating disorder psychopathology. The version of the EDE that was used in this study (version 15.0) is shown in Appendix 3.1.

Psychometric studies of the EDE have shown high interrater reliability (Cooper & Fairburn, 1987; Grilo, Masheb, Lozano-Blanco & Barry, 2004; Rizvi, Peterson, Crow & Agras, 2000; Rosen, Vara, Wendt & Leitenberg, 1990; Wilson & Smith, 1989), high internal consistency (Beumont, Kopec-Schrader, Talbot & Touyz, 1993; Cooper, Cooper & Fairburn, 1989), and good sensitivity to change (Sysko, Walsh & Fairburn, 2005).

To date, the test-retest reliability of the EDE has been reported only twice, both times in studies with a small sample size (Rizvi et al., 2000; Grilo et al., 2004). The findings of

these two studies suggest that the EDE demonstrates good retest-reliability. Clearly more studies with larger sample sizes are needed.

Despite its widespread use, the EDE has undergone relatively limited investigation of various aspects of its validity. This is not unlike the case for other structured interviews for psychiatric problems, given the complexity of designing true validity studies (Malgady, Rogler & Tryon, 1992) and the lack of a clear external validator or biological marker (Wilson, 1993). The EDE has been shown to have good discriminant validity for distinguishing between patients with eating disorders and healthy controls (Cooper et al., 1989; Wilson & Smith, 1989). To date, only one factor-analytic study has been reported (Mannucci, Ricca, Di Bernardo, Moretti, Cabras & Rotella, 1998) to test one specific aspect of construct validity, and this raised questions about the structure of the instrument. Two studies have reported modest concurrent validity for the EDE items pertaining to eating behaviours as tested against prospective self-monitoring records (Rosen et al., 1990; Loeb, Pike, Walsh & Wilson, 1994). These studies highlight the need for further psychometric evaluations of the EDE.

The eating disorder features of interest for this study were the presence and frequency of objective bulimic episodes (DSM-IV binge eating); self-induced vomiting and laxative misuse; attempts to restrict eating due to shape and weight concerns (as assessed by the Dietary Restraint subscale of the EDE); the over-evaluation of shape or weight (as assessed by the Shape Concern and the Weight Concern subscales) and other eating-related features (as assessed by the Eating Concern subscale). Also of interest was the overall severity of the eating disorder as assessed with the Global EDE score. Patients were

weighed and their height was measured at assessment, so that their body mass index (BMI)[3] could be calculated.

Patients' EDE data were used to assign DSM-IV eating disorder diagnoses. Tables 3.1 and 3.2 on the next two pages show the DSM-IV criteria for anorexia nervosa and bulimia nervosa respectively as well as their EDE-based operational definitions. As specified by DSM-IV those eating disorders that did not meet these two operational definitions were classed as cases of eating disorder NOS. In addition, diagnoses of "binge eating disorder" were made based on the research criteria specified in DSM-IV (see Table 1.4, Chapter One). The "Binge eating disorder module" of the EDE (see Appendix 3.1) was used to assess whether the research criteria for binge eating disorder were fulfilled. Binge eating disorder is a provisional new eating disorder diagnosis. Technically, it is an example of the type of clinical problem subsumed under the rubric of eating disorder NOS (see Chapter One, Section 1.3.4).

History of the eating disorder

The history of the eating disorder was taken by one of the three senior clinicians using a standard protocol. The onset of the eating disorder was dated from the first period of sustained (i.e., lasting at least three months) dieting, self-induced vomiting or binge eating.

[3] Body mass index = weight (kg) / [height (m)] ^{2}

Table 3.1 *DSM-IV diagnostic criteria for Anorexia Nervosa and their EDE[1]-based operational definitions*

DSM-IV diagnosis	DSM-IV diagnostic criteria	EDE-based operational definitions
Anorexia Nervosa	A. Refusal to maintain one's body weight at or above a minimally normal weight for age and height (e.g., weight loss leading to maintenance of body weight less than 85% of that expected; or failure to make expected weight gain during period of growth, leading to body weight less than 85% of that expected).	▪ A body mass index below or equal to 17.5 ▪ A rating of 1 on the EDE item "Maintained low weight"
	B. Intense fear of gaining weight or becoming fat, even though underweight.	▪ A rating of higher or equal to 4 on the EDE item "Fear of weight gain" in each of the three months prior assessment
	C. Disturbance in the way in one's body weight or shape is experienced, undue influence of body weight or shape on self-evaluation, or denial of the seriousness of the current low body weight.	▪ A rating of higher or equal to 4 on the EDE item "Importance of weight" or "Importance of shape" in each of the three months prior assessment
	D. In postmenarcheal females, amenorrhea, i.e., the absence of at least three consecutive menstrual cycles. (A woman is considered to have amenorrhea if her periods occur only following hormone, e.g., estrogen administration.)	▪ A rating of 0 or 44 on the EDE item menstruation (no periods in the three months before assessment or taking an oral contraceptive)

[1] Eating Disorder Examination (EDE), version 15.0 (Fairburn & Cooper, 1993)

Table 3.2 *DSM-IV diagnostic criteria for Bulimia Nervosa and their EDE[1]-based operational definitions*

DSM-IV diagnosis		DSM-IV diagnostic criteria	EDE-based operational definitions
Bulimia Nervosa	A.	Recurrent episodes of binge eating. (...)	• 24 or more objective bulimic episodes within the three months prior to assessment and 8 or more bulimic episodes in the month prior to assessment
	B.	Recurrent inappropriate compensatory behavior in order to prevent weight gain, such as self-induced vomiting; misuse of laxatives, diuretics, enemas, or other medications; fasting; or excessive exercise.	• 24 or more episodes of either SIV, or laxative misuse or diuretic misuse over past three months (i.e., total number of "purging" occasions over past 3 months higher or equal to 24) and 8 or more purging episodes in the past month OR extreme dietary restriction (i.e., fasting) outside bulimic episodes, as assessed with the EDE Item "Dietary restriction outside bulimic episodes" (rating of 2)
	C.	The binge eating and inappropriate compensatory behaviors both occur, on average, at least twice a week for 3 months.	• See above
	D.	Self-evaluation is unduly influenced by body shape and weight.	• A rating of higher or equal to 4 on the EDE items "Importance of weight" or "Importance of shape" in each of the last three months
	E.	The disturbance does not occur exclusively during episodes of Anorexia nervosa.	• The diagnosis of anorexia nervosa is made if the diagnostic criteria for both bulimia nervosa and anorexia nervosa are fulfilled.

[1] Eating Disorder Examination (EDE), version 15.0 (Fairburn & Cooper, 1993)

General psychiatric features

Brief Symptom Inventory (BSI; Derogatis & Melisaratos, 1983)

This is a shortened version of the SCL-90-R (Symptom Check List-90-Revised; Derogatis, 1975, 1977), a widely used self-report measure that assesses current psychological distress and symptoms in both patient and non-patient populations. The BSI is the 53-item version of the scale. The BSI is used to measure nine primary symptom dimensions (somatization, obsessive-compulsive behavior, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism) and generates three global indices of distress [Global Severity Index (GSI), Positive Symptom Distress Index (PSDI), and Positive Symptom Total (PST)]. Each item of the BSI is rated on a 5-point scale ranging from 0 (not at all) to 4 (extremely). The BSI measures the experience of symptoms in the past seven days including the day the BSI was completed.

Both test-retest and internal consistency reliabilities have been shown to be very good for the primary symptom dimensions of the BSI. In terms of validation, high convergence between BSI scales and similar dimensions of the MMPI (Minnesota Multiphasic Personality Inventory; Hathaway & McKinley, 1943) provide evidence of convergent validity, and factor analytic studies of the internal structure of the scale contribute evidence of construct validity (Derogatis & Melisaratos, 1983).

For the purpose of this study, only the Global Severity Index (GSI), the most widely used global index of the BSI, was used. It is calculated using the sum for the nine symptom dimensions and dividing by the total number of items to which the individual responded. The GSI combines information about the number of symptoms and the intensity of perceived distress. Of the three global indices the GSI is the most sensitive indicator of current distress levels, and the authors suggest that it should be utilised in most instances where a single

summary measure is required (Derogatis & Melisaratos, 1983). According to the authors the index provides

> *"...the most sensitive single indicator of the respondent's distress level, combining information on the number of symptoms and intensity of distress."*
> (Derogatis & Melisaratos, 1983, p.597)

The GSI is a widely used measure of general psychiatric psychopathology. For the present study, a 28-day time frame was adopted for the BSI, in order to make the measure comparable to other measures used in this study. The BSI is shown in Appendix 3.2.

Social Adjustment Scale – Self-Report *(self-report SAS-M*; Cooper, Osborn, Gath & Feggetter, 1982)

This is the UK version of one of the most widely used self-report measures of current social and interpersonal functioning (SAS-SR; Weissman & Bothwell, 1976). It consists of 45 items assessing social and interpersonal functioning across six domains ("Role Areas"): Work (also housework or work as a student), Extended Family, Marital, Parental, Social and Leisure, and Family Unit. Each item is rated on a five point scale ranging from 0 ("Not at all") to 4 ("All of the time"). Patients' ratings on the questionnaire are interpreted in three ways: (i) functioning in each of the six role areas; (ii) four descriptive categories (Performance, Interpersonal Behaviour, Friction and Feelings); (iii) a single score for all items, the Overall Adjustment Score.

For the purpose of this study, only the Overall Adjustment Score was used. It is calculated using the sum for the 45 items dividing by the total number of items to which the individual responded. Lower scores indicate better social functioning. The psychometric properties of the US version (the SAS-SR) have been supported by a number of studies (Weissman & Bothwell, 1976; Weissman, Prusoff, Thompson, Harding & Myers, 1978; Weissman, Olfson, Gameroff, Feder & Fuentes, 2001). The UK version differs from the

original version in the wording, which was changed to make it more familiar and easily understandable to people in the UK. The UK version of the SAS-SR has shown good agreement (r=.80 for the Overall Adjustment Score) with the original SAS, the interview version (SAS; Weissman and Paykel, 1974) (Cooper et al., 1982).

For the present study, a 28-day time frame was adopted for the SAS-M, in order to make the measure comparable to other measures used in this study. The SAS-M is shown in Appendix 3.3.

Rosenberg Self-esteem Scale (RSE; Rosenberg, 1965)

The Rosenberg Self-Esteem Scale is a widely used 10-item self-report measure of global self-esteem. It consists of 10 statements related to overall feelings of self-worth or self-acceptance. The items are answered on a four-point scale ranging from 'strongly agree' (3) to 'strongly disagree' (0). Five items are reverse scored. The total score consists of the sum of the 10 items. Higher scores indicate a higher level of self-esteem. The scale, originally validated in a large sample of high-school students (Rosenberg, 1965) and thereafter used in non-psychiatric and psychiatric adults (Blascovich and Tomaka, 1993), has demonstrated high reliability: test-retest correlations are typically in the range of 0.82 to 0.88 and Cronbach's alpha for various samples are in the range of 0.77 to 0.88 (Blascovich and Tomaka, 1993).

For the present study, a 28-day time frame was adopted for the RSE, in order to make the measure comparable to other measures used in this study. The RSE is shown in Appendix 3.4.

Data analysis

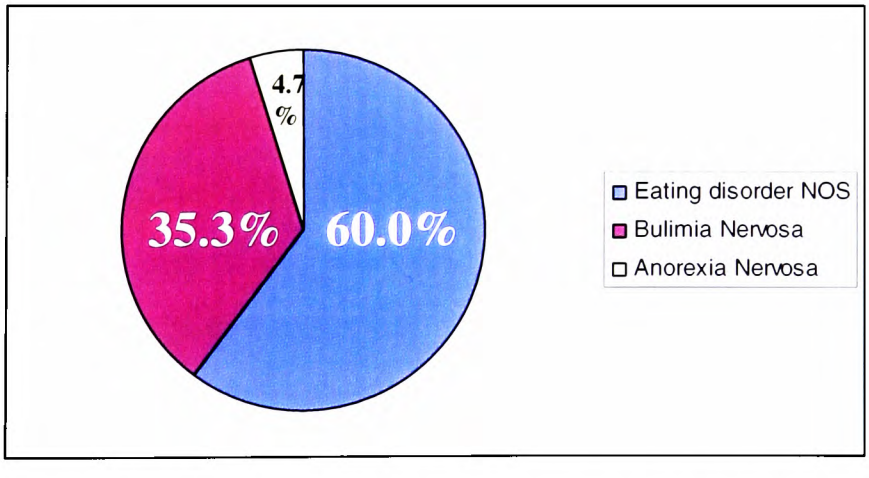
The data are reported as means (SD) for continuous data and n (%) for categorical data. Patients with eating disorder NOS and bulimia nervosa were compared using independent samples t-tests and Mann-Whitney tests for continuous data (normally and non-normally distributed data, respectively) and chi-squared tests for categorical data. Statistical significance was taken at the 5% level ($p<0.05$), with 95% confidence intervals (CIs) used to express the uncertainty around the estimates.

3.2.3 Results

Sample

The sample comprised 170 patients with a clinical eating disorder. One hundred and five of the participants were from Oxford and 65 from Leicester. The operational EDE-based DSM-IV diagnoses of the 170 participants were as expected for adult outpatient eating disorder services and are shown in Figure 3.1.

Figure 3.1 *Relative prevalence rates of the three eating disorders in a sample of 170 adult outpatients*



Eight patients fulfilled diagnostic criteria for anorexia nervosa (4.7%), 60 met diagnostic criteria for bulimia nervosa (35.3%), and 102 were classified as eating disorder NOS (60.0%). Seven patients (4.1%) fulfilled research criteria for binge eating disorder.

The diagnostic distribution at the two centres did not differ significantly. Fifty-six patients in Oxford (53.3%) and 46 patients in Leicester (70.8%) were diagnosed with eating disorder NOS.

Aim 1 – Clinical Characteristics of the Eating Disorder NOS Cases

Table 3.3 shows the characteristics of the full sample and those of the three DSM-IV diagnostic subgroups. The great majority of the eating disorder NOS cases were female, single, white and in their twenties, as were those with anorexia nervosa and bulimia nervosa. Their eating disorder was longstanding, the mean duration being 8.2 years (SD=7.2), and generally its course was unremitting. Almost a quarter (22.5%) had a history of anorexia nervosa and 38.2% had a history of bulimia nervosa. The EDE ratings of the eating disorder NOS cases indicate that they had the eating habits and attitudes to shape and weight that are characteristic of anorexia nervosa and bulimia nervosa. Figure 3.2 on page 77 shows the scores of eating disorder NOS and bulimia nervosa patients on the four EDE subscales as well as their Global EDE score. The scores of a representative sample of young adult women from across Oxfordshire (Beglin, 1990) are also shown in the Figure. It can be seen that the levels of dietary restraint and concerns about shape, weight and eating were markedly elevated in both the bulimia nervosa and the eating disorder NOS cases (for a more detailed comparison see Aim 2 below). In almost three-quarters of the eating disorder NOS cases (73.5%) the overall severity of the eating disorder (Global EDE score) was more than two SDs above the community norm.

Almost half of the eating disorder NOS cases (46.1%) reported recurrent objective bulimic episodes, whereas more than three-quarter (78.4%) showed bulimic episodes of any size (i.e., objective and subjective bulimic episodes). Half (49.0%) reported self-induced vomiting and a quarter (24.5%) misused laxatives. The mean body mass index of the cases

with eating disorder NOS was unremarkable at 22.3 (SD=4.4). Eight (7.8%) had a BMI of 17.5 or below, and eight (7.8%) had a BMI of 30 or more.

The eating disorder NOS cases also had a high level of general psychiatric symptoms and associated distress as shown by their mean BSI Global Severity Index (1.55, SD=0.8). In nearly three-quarters of all eating disorder NOS cases (72.5%) the overall level of general psychiatric symptoms (the GSI) was more than two standard deviations above that found in a random sample of 719 non-patients (0.30, SD=0.3; Derogatis & Melisaratos, 1983). Their psychosocial and interpersonal functioning as well as their level of self-esteem was in the normal range: their mean Overall Adjustment score (SAS-M) (1.54; SD=0.4) and their total score on the RSE (21.3; SD=5.0) did not differ significantly to those found in a community sample (1.59; SD=0.3; Weissman et al., 1978) and in non-psychiatric controls (22.0, SD=4.1; Serretti, Olgiati & Colombo, 2005) respectively.

Nearly a quarter of the eating disorder NOS cases (23.5%) had a weekly alcohol intake above UK recommended maximum levels (14 units/week for women; Austoker, 1994), while only a small proportion misused drugs (2.9%) or engaged in self-harm behaviour (11.8%).

Table 3.3 *The clinical characteristics of the full eating disorder sample and the three diagnostic subgroups*

	Full Eating Disorder Sample (N=170)	**Eating Disorder NOS (N=102)**	Bulimia Nervosa (N=60)	Anorexia Nervosa (N=8)	Comparison of Eating Disorder NOS and Bulimia Nervosa	
					Difference [1] (95% CI)	Test statistic and p-value
Age, years, mean (SD)	25.8 (6.8)	**26.1 (7.2)**	25.5 (6.7)	24.1 (3.4)	0.62 (-1.62, 2.87)	t=0.55, df=160, p=.59
Gender, n (%) female	162 (95.3)	**99 (97.1)**	55 (91.7)	8 (100)	5.4% (-1.7, 15.3)	χ^2 = 1.33, df=1, p=.25
Ethnicity, n (%):						
- White	152 (89.4)	**90 (88.2)**	54 (90.0)	8 (100)	white vs other	χ^2 (white vs other)
- Asian	12 (7.1)	**8 (7.8)**	4 (6.7)	0	-1.8% (-11.1, 9.5)	=0.007, df=1, p=.93
- African British	0	**0**	0	0		
- Mixed	6 (3.5)	**4 (3.9)**	2 (3.3)	0		
Marital status, n (%):						
- single, never married	129 (75.9)	**75 (73.5)**	47 (78.3)	7 (87.5)	single vs other	χ^2 (single vs other)
- married or living as such	38 (22.4)	**26 (25.5)**	11 (18.3)	1 (12.5)	-4.8% (-17.4, 9.4)	0.246, df=1, p=.62
- separated or divorced	3 (1.8)	**1 (1.0)**	2 (3.3)	0		
Occupation, n (%):						
- higher	30 (17.6)	**17 (16.7)**	11 (18.3)	2 (25.0)	student vs other	χ^2 (student vs other)
- intermediate	25 (14.7)	**15 (14.7)**	9 (15.0)	1 (12.5)	0.78% (-14.8, 16.0)	0.00, df=1, p=1.00
- lower	22 (12.9)	**13 (12.7)**	8 (13.3)	1 (12.5)		
- unclassifiable	18 (10.6)	**12 (11.8)**	6 (10.0)	0		
- full-time students	75 (44.1)	**45 (44.1)**	26 (43.3)	4 (50.0)		
Eating disorder psychopathology, mean(SD):						
- overall severity (global EDE[2])	3.35 (1.0)	**3.20 (1.0)**	3.49 (1.1)	4.15 (1.0)	-0.29 (-0.61, 0.039)	M-W z=1.85, p=.065
- above 1SD of community norm[3], n (%)	159 (93.5)	**94 (92.2)**	57 (95.0)	8 (100)	-2.8% (-10.5, 6.7)	χ^2 = 0.14, df=1, p=.71
- above 2SD of community norm[3], n (%)	129 (75.9)	**75 (73.5)**	47 (78.3)	7 (87.5)	-4.8% (-17.4, 9.4)	χ^2 = 0.25, df=1, p=.62
- dietary restraint (EDE subscale[2])	3.50 (1.4)	**3.43 (1.5)**	3.50 (1.3)	4.48 (1.1)	-0.07 (-0.53, 0.39)	M-W z=0.17, p=.87
- eating concern (EDE subscale[2])	2.42 (1.3)	**2.23 (1.2)**	2.58 (1.4)	3.65 (1.3)	-0.35 (-0.77, 0.065)	M-W z=1.49, p=.14
- shape concern (EDE subscale[2])	3.84 (1.2)	**3.64 (1.2)**	4.06 (1.1)	4.58 (1.0)	-0.42 (-0.80, 0.036)	M-W z=2.21, p=.027
- weight concern (EDE subscale[2])	3.63 (1.3)	**3.51 (1.2)**	3.81 (1.4)	3.90 (1.5)	-0.30 (-0.71, 0.11)	M-W z=1.99, p=.047

	Full Eating Disorder Sample (N=170)	**Eating Disorder NOS (N=102)**	Bulimia Nervosa (N=60)	Anorexia Nervosa (N=8)	Comparison of Eating Disorder NOS and Bulimia Nervosa	
					Difference [1] (95% CI)	Test statistic and p-value
Eating disorder behaviour:						
- objective bulimic episodes, n (%) present	108 (63.5)	**47 (46.1)**	60 (100)	1 (12.5)	-53.9%(-63.3, -42.6)	χ^2 = 46.6, df=1, p<.001
if present, episodes/28 days (median)	14.5	**6.0**	22.0	29.0		M-W z=6.54, p<.001
- bulimic episodes of any size, n (%) present	148 (87.1)	**80 (78.4)**	60 (100)	8 (100)	-21.6%(-30.5, -12.4)	χ^2 =13.19, df=1, p<.001
if present, episodes/28 days (median)	22.5	**12.0**	30.0	14.5		M-W z=5.91, p<.001
- self-induced vomiting, n (%) present	104 (61.2)	**50 (49.0)**	52 (86.7)	2 (25.0)	-37.7%(-49.1, -23.2)	χ^2 =21.37, df=1, p<.001
if present, episodes/28 days (median)	20.5	**8.0**	30.0	37.5		M-W z=5.18, p<.001
- laxative misuse, n (%) present	43 (25.3)	**25 (24.5)**	14 (23.3)	4 (50.0)	1.2% (-13.0, 14.0)	χ^2 = 0.00, df=1, p=1.00
if present, episodes/28 days (median)	10.0	**7.0**	20.5	11.0		M-W z=1.92, p=.054
Body mass index, mean (SD)	22.3 (4.4)	**22.3 (4.3)**	23.0 (4.3)	16.7 (0.6)	-0.82 (-2.26, 0.62)	t=1.12, df=160, p=.26
Duration of eating disorder, years, mean (SD)	8.3 (7.0)	**8.2 (7.2)**	9.0 (6.8)	3.8 (2.8)	-0.87 (-3.13, 1.40)	t=0.76, df=160, p=.45
Lowest adult body mass index, mean (SD)	18.4 (2.9)	**18.3 (2.9)**	18.9 (2.9)	16.3 (0.9)	-0.60 (-1.57, 0.36)	t=1.24, df=149, p=.22
Highest adult body mass index, mean (SD)	25.9 (4.9)	**25.9 (4.7)**	26.4 (5.4)	21.8 (1.8)	-0.49 (-2.09, 1.10)	t=0.61, df=158, p=.54
General psychiatric features, BSI[4], mean (SD)	1.61 (0.8)	**1.55 (0.8)**	1.62 (0.8)	2.25 (0.8)	-0.08 (-0.33, 0.18)	t=0.58, df=156, p=.57
Self-esteem, RSE[5], mean (SD)	21.1 (5.5)	**21.3 (5.0)**	21.1 (6.0)	17.0 (5.4)	0.21 (-1.72, 2.14)	t=0.22, df=127, p=.83
Social adjustment, SAS[6], mean (SD)	1.57 (0.5)	**1.54 (0.4)**	1.59 (0.5)	1.87 (0.4)	-0.055 (-0.21, 0.10)	t=0.68, df=128, p=.50
Alcohol intake, units/week, mean (SD)	10.9 (15.1)	**9.6 (15.5)**	14.1 (14.7)	3.4 (2.5)	-4.43 (-9.32, 0.46)	M-W z=2.54, p=.011
Alcohol intake, n (%) with mean weekly intake above recommended maximum level	44 (25.9)	**24 (23.5)**	20 (33.3)	0	-9.8% (-24.3, 4.17)	χ^2 =1.37, df=1, p=.24
Drug misuse, current, n (%)	3 (1.8)	**3 (2.9)**	0	0	-2.9% (-3.4, 8.3)	χ^2 = 0.54, df=1, p=.46
Self-harm, current, n (%)	25 (14.7)	**12 (11.8)**	11 (18.3)	2 (25.0)	-6.6% (-19.2, 4.36)	χ^2 =0.85, df=1, p=.36

[1] Eating disorder NOS score minus bulimia nervosa score

[2] Eating Disorder Examination 15.0 (Fairburn & Cooper, 1993)

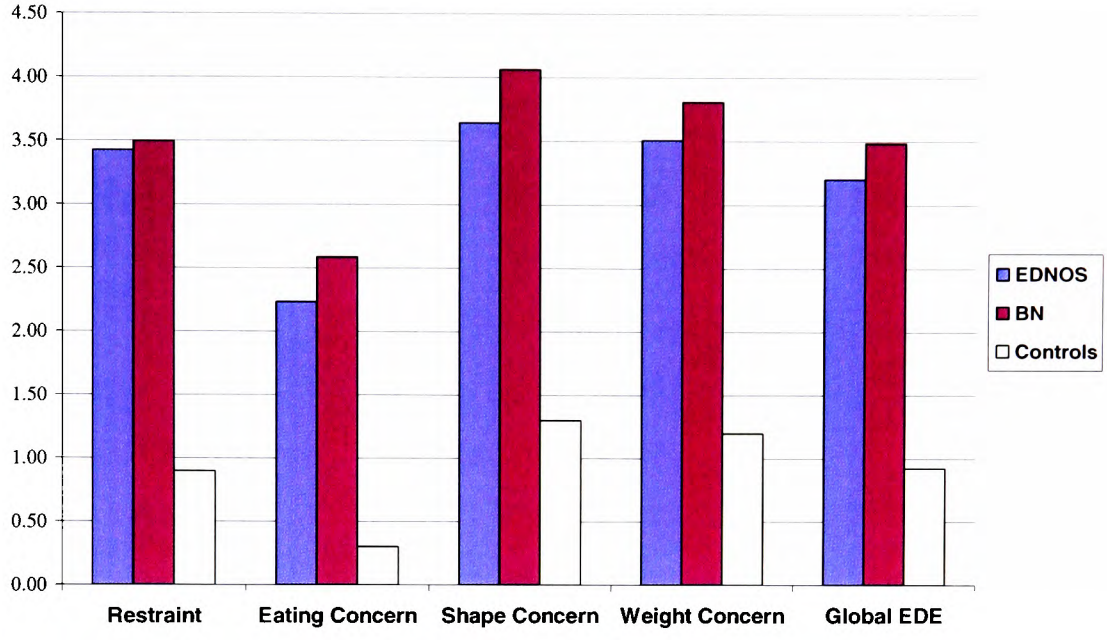
[3] Community EDE norm for young adult women (Beglin, 1990)

[4] Brief Symptom Inventory (Derogatis & Melisaratos, 1983)

[5] Rosenberg Self-Esteem Scale (Rosenberg, 1965)

[6] Social-Adjustment Scale – UK version (Cooper et al., 1982)

Figure 3.2 *Scores of patients with eating disorder NOS (N=102), bulimia nervosa (N=60), and a community sample[1] (N=243) on the four EDE[2] subscales and the Global EDE score*



[1] Beglin, 1990
[2] Eating Disorder Examination (Fairburn & Cooper, 1993)

Table 3.4 below shows key demographic and clinical characteristics of the eating disorder NOS cases at the two centres. The Oxford patients were remarkably similar to the Leicester patients in terms of their age, gender, marital status, the overall severity of their eating disorder psychopathology and the duration of their eating disorder. The Oxford and Leicester groups differed significantly with regard to their BMI, their ethnicity and their level of general psychiatric features as measured with the GSI. The Leicester patients had a higher BMI than Oxford patients (23.3; SD=4.55 vs 21.5; SD=4.10; t=2.1, df=100, p=.038). There was a greater ethnic diversity among the Leicester patients and their average GSI was higher than that of the patients in Oxford (1.76; SD=0.83 vs.1.37; SD=0.72; t=2.5, df=97, p=.015).

Table 3.4 *Demographic and clinical characteristics of eating disorder NOS patients at Oxford and Leicester*

	All eating disorder NOS cases (N=102)	Oxford (N=56)	Leicester (N=46)	Comparison of Oxford and Leicester eating disorder NOS cases
				Test statistic and value
Age, years, mean (SD)	26.1 (7.2)	26.6 (7.32)	25.5 (7.00)	t=0.79, df=100, p=.43
Gender, n (%) female	99 (97.1)	56 (100)	43 (93.5)	χ^2 =1.83, df=1, p=.18
Ethnicity, n (%):				
White	90 (88.2)	56 (100)	34 (73.0)	χ^2 (white vs other)
Asian	8 (7.8)	0	8 (17.4)	= 14.14, df=1, p<.001
African British	0	0	0	
Mixed	4 (3.9)	0	4 (8.7)	
Marital status, n (%)				
- single, never married	75 (73.5)	41 (73.2)	34 (73.9)	χ^2 (single vs other)
- married or living as such	26 (25.5)	14 (25.0)	12 (26.1)	= 0.00, df=1, p=1.00
- separated or divorced	1 (1.0)	1 (1.8)	0	
Severity of ED psychopathology (global EDE[1] score), mean (SD)	3.20 (1.0)	3.12 (1.05)	3.30 (0.89)	z=-.861, p=.39
Body mass index, mean (SD)	22.3 (4.3)	21.5 (4.10)	23.3 (4.55)	t=2.1, df=100, p=.038
Duration of eating disorder, years, mean (SD)	8.2 (7.2)	8.1 (7.15)	8.2 (7.32)	t=0.1, df=100, p=.93
General psychiatric features, BSI[2], mean (SD)	1.55 (0.8)	1.37 (0.72)	1.76 (0.83)	t=2.5, df=97, p=.015

[1] Eating Disorder Examination (Fairburn & Cooper, 1993)
[2] Brief Symptom Inventory (Derogatis & Melisaratos, 1983)

Aim 2 – Comparisons of the Cases of Eating Disorder NOS with Cases of Bulimia Nervosa

The cases of eating disorder NOS were compared with those with bulimia nervosa with respect to their demographic features, the full range of their psychopathology, their eating disorder behaviour and the duration of their eating disorder (see Table 3.3). Certain differences in their eating disorder behaviour were inevitable given the definitions of the two diagnoses.

Demographic data

No significant differences were found in age, gender, marital status, ethnicity or occupation between the eating disorder NOS and the bulimia nervosa patient groups. The majority of patients in both diagnostic groups were in their mid-twenties, female, single and white. In both groups nearly half of the patients (44.1% and 43.3% respectively) were full-time students.

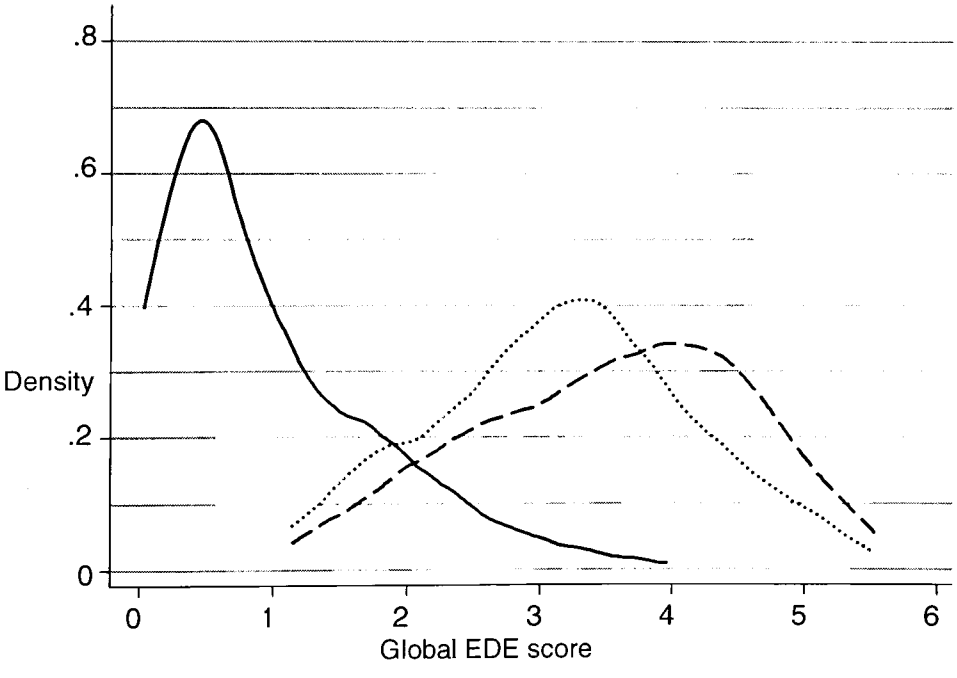
Eating disorder features

As shown above in Table 3.3 and Figure 3.2 above, the two groups were remarkably similar with regard to the overall severity of their eating disorder psychopathology (as measured by the Global EDE score) and their level of dietary restraint and eating concerns (as measured by the Dietary Restraint and Eating Concern subscales of the EDE). A statistically significant difference between the two groups was found on the Shape and Weight Concern subscales of the EDE. Patients with bulimia nervosa scored on average significantly higher on these subscales than patients with eating disorder NOS. As shown in Table 3.3 the 95% CIs for the differences in Shape and Weight Concern subscale scores indicate that the likely true differences between the groups were modest in clinical terms (at most less than one rating point on the EDE). Figure 3.3 shows the distribution of the global EDE scores of the bulimia nervosa and eating disorder NOS cases, together with that of the community sample mentioned above (Beglin, 1990). The similarity of the two patient groups is striking as is their difference from the normative sample.

As would be expected by definition, significantly more patients with bulimia nervosa reported binge eating and self-induced vomiting in the 28 days before the time of assessment than those with eating disorder NOS. Similarly, the frequency of these two forms of behaviour (in those where present) was significantly higher in the patients with bulimia

nervosa. This is not surprising given the DSM-IV twice-weekly frequency threshold for binge eating and compensatory behaviours that is required to make the diagnosis of bulimia nervosa. No significant differences were found between the two groups in the occurrence and frequency of misuse of laxatives for the purpose of controlling weight or shape. Similarly, the body mass index of the two groups did not differ significantly. The average body mass index of both groups was in the middle of a healthy weight range[4]. There was also no significant difference in the mean duration of the eating disorder between patients with eating disorder NOS and those with bulimia nervosa. Both groups had had their disorder for an equally long time, the average being about eight years.

Figure 3.3 *Severity of eating disorder features (global EDE[1] score) in cases of eating disorder NOS (dotted line), cases of bulimia nervosa (dashed line) and in a normative sample of young adult women[2] (normal line)*



[1] Eating Disorder Examination (EDE; Fairburn & Cooper, 1993)
[2] Beglin, 1990

[4] According to the World Health Organization, the healthy weight range is between a body mass index of 18.5 and 25.0 (WHO, 2002).

General psychiatric symptoms

The level of associated general psychopathology as measured by the GSI of the BSI was similarly high for the two patient groups. No significant differences were found between patients with eating disorder NOS and bulimia nervosa with regard to their psychosocial functioning and their level of self-esteem (as measured by the SAS-M and the RSE), both of which were in the normal range for both patient groups. The patients with bulimia nervosa had a significantly higher weekly alcohol intake (measured in units) than those with eating disorder NOS, but the proportion of participants with a weekly alcohol intake above the recommended maximum level did not differ significantly. There was very little drug misuse and a low level of self-harm in both patient groups.

3.2.4 Summary of findings

- In a representative sample of 170 adult outpatients with a clinical eating disorder the majority of cases (60%) were classified as eating disorder NOS. A third of the sample (35.3%) met diagnostic criteria for bulimia nervosa, and 4.7% fulfilled diagnostic criteria for anorexia nervosa.
- The eating disorder NOS cases were found to have eating disorder features of the type and severity of those seen in patients with anorexia nervosa and bulimia nervosa: they presented with markedly elevated levels of dietary restraint and concerns about eating, weight and shape. Half of them presented with objective bulimic episodes and self-induced vomiting. A quarter of the eating disorder NOS cases misused laxatives. Their level of general psychiatric symptoms and associated distress was high. Their mean BMI was unremarkable at 22.3 (SD=4.4)
- The eating disorder NOS cases were remarkably similar to those with bulimia nervosa with regard to their demographic features, the full range of their

psychopathology and the duration of their eating disorder. The only statistically significant differences between the two groups were:

A higher frequency of objective bulimic episodes and self-induced vomiting in cases with bulimia nervosa (as would be expected, given the definition of the diagnosis)

Marginally higher scores on the Shape and Weight Concern subscales of the EDE in the cases with bulimia nervosa

A higher weekly alcohol intake in the cases with bulimia nervosa.

- The eating disorder cases in Oxford were remarkably similar to those in Leicester in terms of their age, gender, marital status, the overall severity of their eating disorder psychopathology and the duration of their eating disorder. They differed with regard to the body mass index (higher in Leicester patients), their ethnicity (more diverse in Leicester) and their level of general psychiatric features higher in Leicester patients).

3.3. Characterising Eating Disorder NOS further

3.3.1 Introduction and Aims

As discussed in Chapter One (Section 1.4.2) only few studies exist that have tried to characterise the eating disorder features of people with eating disorder NOS in detail and define subgroups within this diagnostic category. Most of these studies have important limitations; the two most common ones being the use of a community sample (and therefore the absence of an assessment of impairment) and the failure to study the whole group of eating disorder NOS cases (and subgroups instead) (see Chapter One, Section 1.4.2). The present study had two aims. The first was to characterise in detail the eating disorder features of a representative sample of patients with eating disorder NOS by examining the occurrence and frequency of four key eating disorder behaviours. The second was to

identify the most common clinically significant behavioural subgroups within this diagnostic category.

3.3.2 Methods

Design

The sample consisted of the 102 patients with eating disorder NOS described above. Details of the sample and how it was recruited and assessed are provided in section 3.2.2.

Aim 1: In order to describe the eating disorder features of these patients in detail, the occurrence and frequency of four key eating disorder behaviours over the previous 28 days was identified, as assessed using the EDE. The four behaviours were:

- episodes of binge eating (objective bulimic episodes on the EDE)
- episodes of "purging" (defined here as episodes of self-induced vomiting or episodes of laxative misuse)
- driven exercising
- undereating, an index of which is the active maintenance of an unhealthily low weight.

Aim 2: In order to address the second aim, the occurrence of each of these four eating disorder behaviours at a clinically significant level was described both in isolation and in combination. In a next step the most common subgroups were determined.

Clinically significant levels for the four eating disorder behaviours were defined as follows:

- *Binge Eating*: A minimum frequency of four objective bulimic episodes over the previous 28 days (equivalent to an average frequency of once per week). This threshold is in line with recommendations for adjusting the current twice-weekly threshold for binge eating used by DSM-IV. It is based upon evidence that patients

who binge eat once weekly on average do not differ meaningfully from patients who binge eat twice weekly or more (see Chapter One, Section 1.3.5.2).

- *Purging:* A minimum frequency of four episodes over the previous 28 days (equivalent to an average frequency of once per week). This definition follows the convention of using the same frequency criterion for purging as for binge eating.
- *Driven exercising*: The definition of a clinically significant (i.e., pathological) level of driven exercising was decided after consultation with Professor Fairburn, who has been responsible for the successive editions of the EDE. Pathological exercising is difficult to define and discriminate from a high level of "healthy" exercising (as reflected in the absence of a 'threshold' for driven exercising; see Chapter One, Section 1.3.5.2). The "Driven exercising" item of the EDE (version 15.0; see Appendix 3.1) is designed to identify the key psychopathological characteristics of such exercising (the drive experienced to engage in it and the salience of the behaviour). For the purposes of the present study the level of such exercising taken as being "clinically significant" was purposefully set high to avoid falsely classifying intense healthy exercising as pathological. The specific level chosen was based on the clinical features of patients who exercise in this way. It was engaging in driven exercising (as defined by the EDE) for at least 1,200 minutes over the previous month, a threshold equivalent to exercising in a driven way on five days a week for at least one hour.
- *Undereating:* The index of clinically significant undereating was taken as actively maintaining a low body weight at a level that is associated with health risks. The World Health Organisation's (WHO) lower threshold for a healthy weight, a body mass index of 18.5, was adopted (WHO, 2002). Eating disorder NOS cases who

were actively maintaining a body weight below this threshold were regarded as showing a clinically significant level of undereating.

Participants

The sample for this study consisted of the 102 patients with eating disorder NOS that were described in section 3.2.

Measures

As described in section 3.2, patients were assessed using the EDE, version 15.0 (Fairburn & Cooper, 1993). The EDE items of interest for this study were the presence and frequency of objective bulimic episodes, self-induced vomiting, laxative abuse, and driven exercising in the 28 days prior to assessment. The patient's weight and height was used to calculate the body mass index (BMI). The body mass index was used to a) examine the BMI distribution of the sample and b) to determine the number of patients with a BMI below 18.5. The EDE item "Maintained low weight" (EDE version 15.0; see Appendix 3.1) was used to determine whether the participants with a BMI below 18.5 were actively maintaining their weight at this low level.

Data analysis

The data are reported in terms of the number and percentage of participants engaging in the four eating disorder behaviours and the mean frequency (and SD) of these behaviours.

3.3.3 Results

Aim 1 Characterising the eating disorder features in detail

Table 3.5 shows the presence and frequency of the four eating disorder behaviours in the 102 eating disorder NOS cases.

Table 3.5 *Presence and Frequency of the four eating disorder behaviours in patients with eating disorder NOS*

Behaviour	N (%) of eating disorder NOS cases who show feature	Mean frequency (SD; range) in past 28 days in those where present
Binge eating[1]	47 (46.1)	8.9 (9.3; 1-36)
Purging	58 (56.9)	19.3 (19.7; 1-84)
Driven exercising	44 (43.1)	883.1 mins (887.3; 75-4200)
Undereating (BMI < 18.5)	20 (19.6)	17.6[2] (0.65; 16.46-18.48)

[1] Binge eating is defined here as objective bulimic episodes
[2] Mean body mass index in those < 18.5

Binge eating

Nearly half of the eating disorder NOS cases (46.1%) had engaged in binge eating (i.e., experienced at least one objective bulimic episode) over the 28 days before the time of assessment. The mean frequency of binge eating in these patients was 8.9 (SD=9.3) episodes in the last 28 days.

Purging (Episodes of self-induced vomiting or episodes of laxative abuse)

Fifty-eight of the 102 eating disorder NOS cases (56.9%) had engaged in episodes of purging at least once in the previous 28 days. More than half of these 58 patients (56.9%) engaged in self-induced vomiting but no laxative misuse, eight (13.8%) had only misused laxatives, and just over a quarter (29.3%) had engaged in both self-induced vomiting and laxative misuse. The mean frequency of episodes of purging in the 58 patients was 19.3 (SD=19.7).

Driven exercising

Forty-four of the eating disorder NOS cases (43.1%) had exercised in a driven way over the 28 days before assessment. The average number of times that they had engaged in driven exercising was 12.3 (SD=8.0; median=11.0; range=1-28), i.e., on average three times per week. The average length of each episode was 75.2 minutes (SD=51.7; median=60

mins; range=15-300 mins). The eating disorder NOS cases had on average spent a total of 883.1 minutes (SD=887.3) or 14.7 hours engaged in driven exercising over the previous 28 days and an average of 221.84 minutes (3.7 hours) per week.

BMI distribution / Maintenance of an unhealthily low weight

Twenty of the 102 eating disorder NOS patients (19.6%) had a BMI below 18.5. All of these patients were actively maintaining their weight at this low level. Fifty-eight (56.9%) were at a healthy weight with a BMI between 18.5 and 24.9, 16 patients (15.7%) were overweight with a BMI between 25.0 and 29.9, and eight (7.8%) had a BMI above 30.0[5].

Aim 2 Presence of clinically significant levels of the four eating disorder behaviours

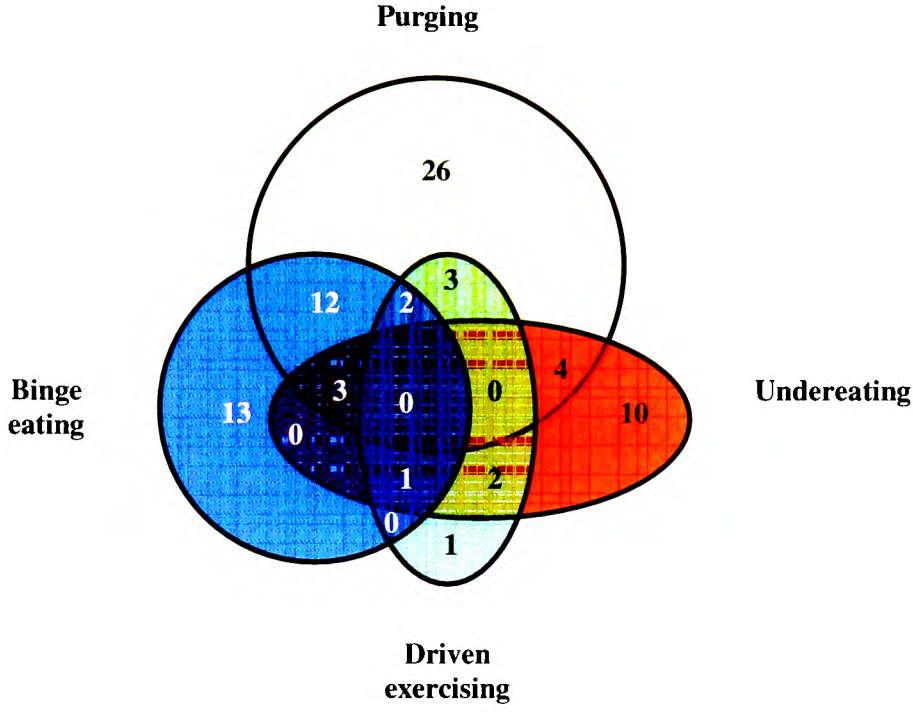
The number and percentage of eating disorder NOS patients who reported clinically significant levels of the four eating disorder behaviours over the 28 days before assessment were as follows:

- Binge eating on average once per week - 31 (30.4%),
- Purging on average once per week - 50 (49.0%),
- Driven exercising for on average 5 hours per week - 9 (8.8%),
- Undereating (actively maintaining a BMI below 18.5) - 20 (19.6%).

The inter-relationship between the four behaviours of interest is shown in the Venn diagram in Figure 3.4. This shows the number of people who reported the four eating disorder behaviours at a clinically significant level, both in isolation and in combination with each other.

[5] The WHO weight guidelines for adults are: BMI 18.5 to 24.9: healthy/normal; BMI 24.9 to 29.9: pre-obese; BMI 30.0 to 39.9: obese; BMI 40 and more: very obese (WHO, 2002)

Figure 3.4 *Venn diagram showing number of eating disorder NOS patients with clinically significant levels[1] of one or more of four eating disorder features over the past month*
(NB: Binge eating is defined here as objective bulimic episodes only)



[1] Here defined as: for binge eating (here OBEs) ≥ once per week
for purging: ≥ once per week
for undereating: BMI < 18.5
for driven exercising: ≥ 300 minutes per week

There are several points to note with regard to Figure 3.4:

1. Seventy-seven (75.5%) of the 102 patients with eating disorder NOS presented with a clinically significant level of at least one of the four eating disorder behaviours and are therefore included in Figure 3.4.

2. Four subgroups account for nearly all the cases in this sub-sample of eating disorder NOS patients (76/77; 98.7%):

   Group I - Twenty-nine patients (28.4% of the full eating disorder NOS sample) purged but were not underweight and did not engage in binge eating

   Group II - Twenty patients (19.6% of the full eating disorder NOS sample) engaged in under-eating (i.e., were underweight)

   Group III - Fourteen patients (13.7% of the full eating disorder NOS sample) reported binge eating and purging but were not underweight

   Group IV Thirteen patients (12.7% of the full eating disorder NOS sample) reported binge eating but were not underweight and did not engage in the other forms of behaviour.

The main clinical characteristics of these four groups are shown in Table 3.6.

Table 3.6 *Clinical characteristics of the four main subgroups of patients with eating disorder NOS who showed a clinically significant level of one or more of four eating disorder behaviours* (NB: Binge eating is defined here as objective bulimic episodes.)

	Group I "Purging" N=29	Group II "Under-eating/ underweight" N=20	Group III "Binge eating[1] and purging" N=14	Group IV "Binge eating only" N=13
Age, years, mean (SD)	26.3 (7.3)	23.3 (4.3)	23.8 (4.6)	30.4 (7.94)
Severity of ED psychopathology (global EDE[2] score), mean (SD)	3.47 (0.79)	2.80 (0.96)	3.74 (1.10)	2.43 (0.74)
Body mass index, mean (SD)	22.2 (2.8)	17.6 (0.7)	22.8 (2.73)	26.5 (5.26)
Duration of eating disorder, years, mean (SD)	8.6 (7.2)	6.5 (5.7)	5.5 (2.71)	10.4 (8.50)
General psychiatric features, BSI[3], mean (SD)	1.66 (0.77)	1.65 (0.85)	1.71 (0.80)	1.12 (0.52)

[1] Binge eating is defined here as objective bulimic episodes
[2] Eating Disorder Examination 15.0 (Fairburn & Cooper, 1993)
[3] Brief Symptom Inventory (Derogatis & Melisaratos, 1983)

As can be seen in Table 3.6, groups I and III resemble each other in several ways: they present with a high level of both eating disorder psychopathology and general psychiatric features, and their BMI is in the mid-healthy range. Group I resembles a state recently referred to as *'purging disorder'* (Keel et al., 2005). Group III could be described as having *'subthreshold bulimia nervosa'*. Group II could be described as *'subthreshold anorexia nervosa'* with a mean BMI of 17.6 and a lower level of eating disorder psychopathology than the two previous groups, which is typical of patients with anorexia nervosa (Couturier & Lock, 2006). Group IV resembles *'binge eating disorder'* in that patients are older, they have a lower level of eating disorder psychopathology, and their mean BMI is above the healthy range (Grilo, 2006).

Twenty-five of the 102 patients (24.5%) are not included in the diagram, as they did not present with clinically significant levels of the four features of interest. Nine of them were at a low healthy weight (BMI 18.5 to 20.0) and engaged in driven exercising at a level that was not severe enough to be classed as clinically significant (as defined above). The remaining 16 presented with objective bulimic episodes and/or episodes of purging at a sub-clinical level. The clinical characteristics of the 25 patients that were not included in the Venn diagram were as follows: the mean age was 27.3 (SD=8.5) and they had had their eating disorder for on average 9.4 years (SD=8.8). The overall severity of their eating disorder psychopathology was in the high range with a mean Global EDE score of 3.31 (SD=0.97) and their mean BMI was 23.7 (SD=8.5). The associated level of general psychopathology was low with a mean GSI of 1.47 (SD=0.87).

3.3.3.1 Sub-analysis

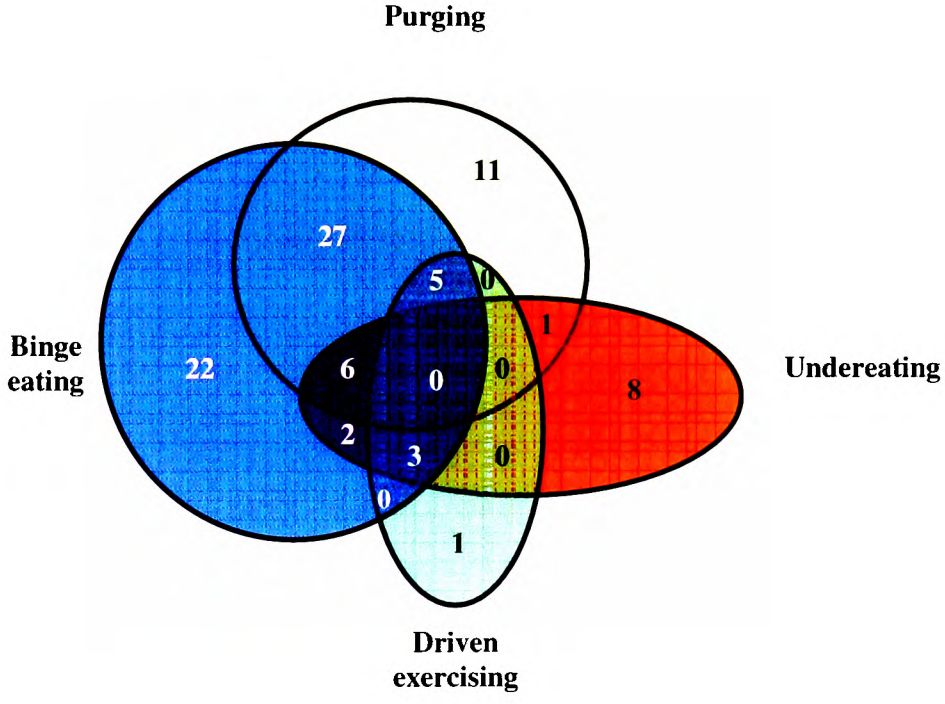
In view of the findings of the previous study (especially the fact that the largest subgroup of patients within eating disorder NOS that showed a 'clinically significant' level

of one or more of the four key behaviours was the one that reported purging but no binge eating), it was decided to undertake a further analysis. This was to examine the occurrence and frequency of binge eating in this subsample of eating disorder NOS cases with the definition of binge eating being expanded to include subjective as well as objective bulimic episodes (i.e., the requirement that binges be large was waived). The relationship between the occurrence of this form of behaviour and the presence of the three other eating disorder behaviours (i.e., purging, driven exercise, and under-eating) at a clinically significant level was also examined.

Results

Eighty of the 102 patients with eating disorder NOS (78.4%) reported bulimic episodes of any size over the 28 days before the time of assessment. The mean frequency of binge eating of this type was 19.4 (SD=21.6) episodes in the last 28 days. Nearly two-thirds (65/102; 63.7%) had shown bulimic episodes of any size at the "clinically significant" frequency of at least four episodes over the past 28 days. The co-occurrence of this behaviour with the other three behaviours is shown in the Venn diagram in Figure 3.5.

Figure 3.5 *Venn diagram showing number of eating disorder NOS patients with clinically significant levels[1] of one or more of four eating disorder features over the past month*
(NB: Binge eating is defined here as objective AND subjective bulimic episodes.)



[1] Here defined as: for binge eating (here: OBEs and SBEs): ≥ once per week
for purging: ≥ once per week
for undereating: BMI < 18.5
for driven exercising: ≥ 300 minutes per week

As can be seen the expanded definition of binge eating resulted in a substantial change to the sub-grouping within eating disorder NOS. Several points are of note:

1. Eighty-six (84.3%) of the 102 patients with eating disorder NOS reported a clinically significant level of at least one of the four eating disorder behaviours and were therefore eligible for inclusion in the Venn diagram.
2. The following four subgroups accounted for nearly all these cases (85/86; 98.8%):

   Group I – Thirty-two patients (31.4% of the full eating disorder NOS sample) reported binge eating and purging but were not underweight

   Group II – Twenty-two patients (21.6% of the full eating disorder NOS sample) reported binge eating but were not underweight and did not engage in the other forms of behaviour

   Group III Twenty patients (19.6% of the full eating disorder NOS sample) engaged in under-eating (i.e., were underweight)

   Group IV – Eleven patients (10.8% of the full eating disorder NOS sample) purged but were not underweight and did not engage in the other forms of behaviour.

Table 3.7 *Clinical characteristics of the four main subgroups of patients with eating disorder NOS who showed a clinically significant level of one or more of four eating disorder behaviours* (NB: Binge eating is defined here as bulimic episodes of any size)

	Group I "Binge eating[1] and purging" N=32	Group II "Binge eating[1] only" N=22	Group III "Under-eating/ underweight" N=20	Group IV "Purging only" N=11
Age, years, mean (SD)	25.0 (6.7)	30.3 (0.8)	23.3 (4.3)	26.9 (6.5)
Severity of ED psychopathology (global EDE[2] score), mean (SD)	3.69 (0.90)	2.75 (0.83)	2.80 (0.96)	3.18 (0.74)
Body mass index, mean (SD)	21.8 (2.6)	26.4 (5.1)	17.6 (0.7)	24.0 (3.0)
Duration of eating disorder, years, mean (SD)	8.2 (6.8)	11.7 (8.3)	6.5 (5.7)	5.8 (4.0)
General psychiatric features, BSI[3], mean (SD)	1.81 (0.73)	1.21 (0.61)	1.65 (0.85)	1.28 (0.78)

[1] Binge eating is defined here as bulimic episodes of any size
[2] Eating Disorder Examination 15.0 (EDE; Fairburn & Cooper, 1993)
[3] Brief Symptom Inventory (Derogatis & Melisaratos, 1983)

The main clinical characteristics of Groups I to IV are shown in Table 3.7.

As can be seen in Table 3.7, the four subgroups had similar characteristics to those in the first analysis, in which binge eating was defined conventionally (see Table 3.6). The number of patients in each group, however, changed substantially: the most common group was now the one resembling *subthreshold bulimia nervosa* (Group I). The group that resembled *purging disorder* (Group IV) was now the least common subgroup (N=11).

With the wider definition of binge eating being used, only 16 patients (15.7% of the full eating disorder NOS sample) were not eligible for inclusion in the Venn diagram, as they did not present with clinically significant levels of any of the four behaviours. Their clinical characteristics were as follows: the mean age was 25.8 (SD=8.2) and they had had their eating disorder for on average 7.4 years (SD=8.6). The overall severity of their eating disorder psychopathology was in the high range with a mean Global EDE score of 3.26 (SD=1.03) and their mean BMI was 22.5 (SD=3.6). The associated level of general psychopathology was low with a mean GSI of 1.41 (SD=0.83).

3.3.4 Summary of findings

- Nearly the entire sample of eating disorder NOS cases (93.1%) reported one or more of the four key eating disorder behaviours (at any level) over the month before assessment.
- Binge eating (i.e., objective bulimic episodes) was present in 46.1% of the cases, purging in 56.9%, driven exercising in 43.1% and undereating in 20.0%.
- Seventy-seven (75.5%) of the 102 cases with eating disorder NOS presented with a clinically significant level of at least one of the four eating disorder behaviours. Amongst these, four subgroups accounted for nearly all of the cases. The first (28.4% of the full eating disorder NOS sample) was a group that purged but was not

underweight and did not engage in any of the other behaviours, and therefore resembled *"purging disorder"* (Keel et al., 2005). The second (19.6%) engaged in clinically significant under-eating (i.e., was underweight) and resembled *subthreshold anorexia nervosa.* The third (13.7%) reported binge eating and purging but was not underweight. These cases might be described as having *subthreshold bulimia nervosa.* The fourth group (12.7%) reported only binge eating but was not underweight and resembled *binge eating disorder.*

- Broadening the definition of binge eating to include bulimic episodes of any size led to an increase in the number of eating disorder NOS patients presenting with a clinically significant level of one or more of the four behaviours (from 75.5% to 84.3%). It also changed the relative prevalence of the four above mentioned subgroups. It increased the *'subthreshold bulimia nervosa'* group from 13.7% to 31.4% (of the full eating disorder NOS sample), and decreased the *'purging disorder'* group from 28.4% to 10.8%.

3.4 Discussion

In this chapter, two studies are described. Their findings are summarised in turn and then discussed together. General conclusions, based on the findings of both studies, are drawn at the end.

Study One

The first study had two aims. The first was to describe the clinical characteristics of adult outpatients with the DSM-IV diagnosis eating disorder NOS by assessing a large and representative patient sample using standardised instruments. The second was to establish the severity of eating disorder NOS by comparing these cases with those with bulimia nervosa, the second most common eating disorder. The sample used for this study was a

sample of adult outpatients with clinical eating disorders recruited at two centres in the UK for a transdiagnostic treatment trial. The diagnostic composition of the sample was as expected with over half the patients receiving the diagnosis eating disorder NOS, a third meeting diagnostic criteria for bulimia nervosa and the remainder having anorexia nervosa. These figures correspond to those from five other studies of representative, well-diagnosed outpatient samples of adult patients with an eating disorder (Martin et al., 2000; Millar et al., 1998; Nollett & Button, 2005; Ricca et al., 2001; Turner & Bryant-Waugh, 2004).

The patients with eating disorder NOS had longstanding eating problems, the mean duration being over eight years. Their EDE ratings showed that they displayed the psychopathology that is characteristic of anorexia nervosa and bulimia nervosa: high dietary restraint and extreme concerns about shape, weight and eating. In almost three-quarters of the eating disorder NOS patients the overall severity of the eating disorder was more than two SDs above the community norm for young women their age. Almost half reported recurrent binge eating and self-induced vomiting. Their mean BMI was unremarkable at 22.3. Almost a quarter of the patients had a history of anorexia nervosa and over a third had a history of bulimia nervosa illustrating the cross-diagnostic temporal movement that is common among patients with eating disorders (Fairburn & Harrison, 2003). Their level of comorbid general psychiatric symptoms was also high. Their psychosocial functioning and level of self-esteem was in the normal range. Nearly a quarter of the eating disorder NOS cases had a weekly alcohol intake above the UK recommended maximum level, while only a small proportion misused drugs or engaged in self-harm behaviour.

There was a similar proportion of eating disorder NOS cases in the Oxford and Leicester samples. The eating disorder NOS patients at the two centres were remarkably similar with regard to their demographic and clinical characteristics, the only statistically

significant differences being a higher BMI, a greater ethnic diversity and a higher level of associated general psychopathology in the Leicester patients.

The eating disorder NOS cases were remarkably similar to those with bulimia nervosa with regard to their demographic features, the full range of their psychopathology and the duration of their eating disorder. There were three statistically significant differences between the two groups: first, there was a higher frequency of objective bulimic episodes and self-induced vomiting in the cases with bulimia nervosa (as to be expected, given the definition of the diagnosis); second, the cases of bulimia nervosa had higher scores on the Shape and Weight Concern subscales of the EDE, although the difference was not significant in clinical terms; and third, the bulimia nervosa cases had a higher weekly alcohol intake that those with eating disorder NOS.

Study Two

The second study reported in this chapter had two aims. The first was to describe in detail the clinical features of patients with eating disorder NOS by examining the occurrence and frequency of four key eating disorder behaviours (binge eating [here defined as objective bulimic episodes], purging, driven exercising and undereating). The second was to identify the most common clinically significant behavioural subgroups within this diagnostic category. The sample used for this study was the sample of eating disorder NOS patients described in Study One of this chapter.

Over ninety percent of the eating disorder NOS cases reported one or more of four eating disorder features (at any level). Nearly half had experienced binge eating over the month before the time of assessment. Thirty percent had shown this eating disorder feature at a clinically significant level (i.e., on average at least once per week).

Just over half of the eating disorder NOS cases reported episodes of purging in the previous month. The majority engaged in self-induced vomiting only, a small proportion of patients only took laxatives, and the remainder engaged in both forms of behaviour. Half the total sample (and therefore almost everyone who engaged in purging) had experienced purging episodes at a clinically significant level.

Over fourty percent of the entire sample of eating disorder NOS cases had exercised in a driven way over the month before assessment. The average frequency was three times per week for an average of about one hour. Less than ten percent of the whole sample had engaged in driven exercising at a level that was defined as clinically significant for the present study.

Twenty percent of the eating disorder NOS cases were actively maintaining their weight at a BMI below 18.5, a weight that is associated with major health risks. This was taken as an index of undereating.

Examining the presence of a 'clinically significant' level of one or more of the four eating disorder behaviours, yielded the following findings: Three quarters of the cases engaged in at least one of the four eating disorder behaviours at a clinically significant level. Amongst these, four subgroups accounted for almost eighty percent of the cases. The first (28.4% of the full eating disorder NOS sample) was a group that purged but was not underweight and did not engage in objective bulimic episodes, and therefore resembled a state recently described as *"purging disorder"* (Keel et al., 2005). The second (19.6% of the full eating disorder NOS sample) was underweight and might be viewed as having *subthreshold anorexia nervosa*. The third subgroup (13.7% of the full eating disorder NOS sample) reported binge eating and purging but was not underweight. These cases might be described as having *subthreshold bulimia nervosa*. The fourth group (12.7% of the full

eating disorder NOS sample) reported only binge eating but was not underweight and resembled *binge eating disorder.*

An additional analysis designed to examine whether the size of binges materially affected this subgrouping showed that the 'purging disorder' group was largely composed of people with subjective bulimic episodes and this group might therefore be better classified within the 'subthreshold bulimia nervosa' group.

Having characterised eating disorder NOS further and made the above distinctions it must be pointed out that the most prominent feature of this sample of eating disorder NOS patients was what they had in common with patients with bulimia nervosa and anorexia nervosa, namely the distinctive psychopathology of patients with eating disorders, which was not examined in this substudy.

Methodological Considerations

The two studies in this chapter had certain strengths. First, the two-site catchment area sampling frame meant that the patients are likely to have been representative of many other outpatient samples of adults with eating disorders. Second, the cases were well characterised and leading measures of psychopathology were employed. Third, although there was multiple statistical testing, it was chosen not to adjust the significance level when comparing the eating disorder NOS cases with those with bulimia nervosa. This favoured the identification of false positives rather than false negatives (i.e., raising Type I rather than Type II error). This was a conservative strategy since it made more likely the detection of statistically significant differences between the groups. If the significance level had been reduced to the 1% level, none of the observed differences would have been statistically significant. The second study in this chapter had an additional strength: it is the only study

to date to have characterised in such detail the occurrence and frequency of key eating disorder behaviours in a representative sample of eating disorder NOS cases.

In terms of limitations, the sample did not include patients at the extremes of the weight spectrum (i.e., those with a BMI <16.0 or ≥40.0). Data from the Leicester clinic for the three-year period 2003 to 2005 indicate that these two group constituted 8% (39/490) and 5% (23/490) respectively of the referrals (Palmer, 2006; personal communication). Had they been included in the present study patients with a BMI below 16 would have been likely to have fulfilled diagnostic criteria for anorexia nervosa whereas those with a BMI of 40 or more would have been likely to have been cases of eating disorder NOS. If these two groups had been included in the sample the diagnostic distribution would have been 11% anorexia nervosa, 31 % bulimia nervosa and 58% eating disorder NOS, figures that differ little from those of the actual study sample. The exclusion of these two small subgroups of patients will have had little effect on the study findings as these concern the characteristics of the eating disorder NOS patients and their severity relative to bulimia nervosa.

One other point about the sample should be noted. The study was of adults who were outpatients with an eating disorder: its findings cannot be generalised to adolescents with an eating disorder. Further studies of other representative samples are needed.

A potential limitation of both studies is the one-month time frame over which the eating disorder features were assessed. It is possible that the 28 days before assessment were not representative of the person's state. A longer time frame (e.g., the past three months) might have provided a more representative picture of the patient's eating disorder although participants could not have been expected to provide the level of detailed information required for the present analyses. A limitation specific to the second study concerns the definitions chosen to define the four behaviours as clinically significant. Each could be

challenged as there is no consensus in this regard. Alternative definitions could be explored. A further point to note with regard to the second study is that only behavioural features were assessed in this study, and not the attitudinal aspects of eating disorder psychopathology (i.e., concerns about shape, weight, and/or controlling eating). If one wanted to know exactly how many of the eating disorder NOS cases could be classed as 'subthreshold anorexia nervosa' or 'subthreshold bulimia nervosa' cases, it would be necessary to also assess the core eating disorder psychopathology. In addition, a three-month time frame would have to be used. This has been done in the second study of Chapter Four (Section 4.3).

Conclusions

The main conclusions to be drawn from these two studies are as follows: First, the findings confirm that eating disorder NOS is the most common eating disorder diagnosis encountered in adult outpatient settings. Second, its psychopathology closely resembles that of anorexia nervosa and bulimia nervosa. Third, across a wide range of clinical variables eating disorder NOS is comparable in severity to bulimia nervosa. Fourth, the majority of patients with eating disorder NOS present with clinically significant levels of one or more key eating disorder behaviours. Fifth, about half of the patients with eating disorder NOS present with eating disorder behaviours that resemble subthreshold bulimia nervosa or subthreshold anorexia nervosa states.

The findings have important practical implications. They highlight the need for research on eating disorders to have a broader focus than at present. Research needs to address the whole range of problems seen in clinical practice, not just anorexia nervosa and bulimia nervosa. Particularly pressing is the need for research on the treatment of eating disorder NOS. Existing treatments for anorexia nervosa and bulimia nervosa might benefit

these patients, given that they share much of the same distinctive psychopathology (Fairburn et al., 2003), but this needs to be shown to be the case.

CHAPTER FOUR

Re-conceptualising Eating Disorder Not Otherwise Specified (Eating Disorder NOS)

4.1 Introduction

Eating disorder not otherwise specified (eating disorder NOS) is the most common eating disorder diagnosis made in most adult outpatient settings (Martin et al., 2000; Millar et al., 1998; Nollett & Button, 2005; Ricca et al., 2001; Turner & Bryant-Waugh, 2004). It is not a mild eating disorder but in severity is comparable to bulimia nervosa (see Chapter Three, Section 3.2). Despite this, it has been neglected in terms of clinical guidance and research. To date no studies have been published on its treatment.

As pointed out in Chapter One, the NOS category in DSM-IV is not intended to be the most common category within a class of disorders. It is meant to be a "category within a class of disorders that is residual to the specific categories in that class..." (APA, 1994). In other words it is meant to denote a small residual group. This is clearly not the case in the eating disorders. The two specified diagnoses anorexia nervosa and bulimia nervosa "do not cover the (eating disorder) ground" (Palmer, 2003). The largest category within the eating disorders is the "residual" one of eating disorder NOS.

Thus there are two problems with the diagnosis. The first is the nosological one and the second is its neglect. This chapter focuses on these problems from a conceptual perspective. In section 4.2 three ways of re-conceptualising the diagnosis eating disorder NOS are proposed and their potential for solving the problems of nosology and neglect associated with the diagnosis are discussed. These proposals are the product of joint work with Professor Christopher Fairburn and were described in a paper published in *Behaviour Research and Therapy* in June 2005 (Fairburn & Bohn, 2005). In section 4.3 the impact of the first of the proposed solutions is examined, namely, relaxing the diagnostic criteria for

anorexia nervosa and bulimia nervosa in a well-diagnosed and representative sample of eating disorder cases. In section 4.4 a study of the clinical utility (as defined by First and colleagues (First, Pincus, Levine, Williams, Ustun & Peele, 2004)) of the second of the three proposals is described. Each study and its findings are briefly summarised at the end of each section and the group of studies is considered in the overall discussion in section 4.5.

4.2 Three solutions to the problem of Eating Disorder NOS

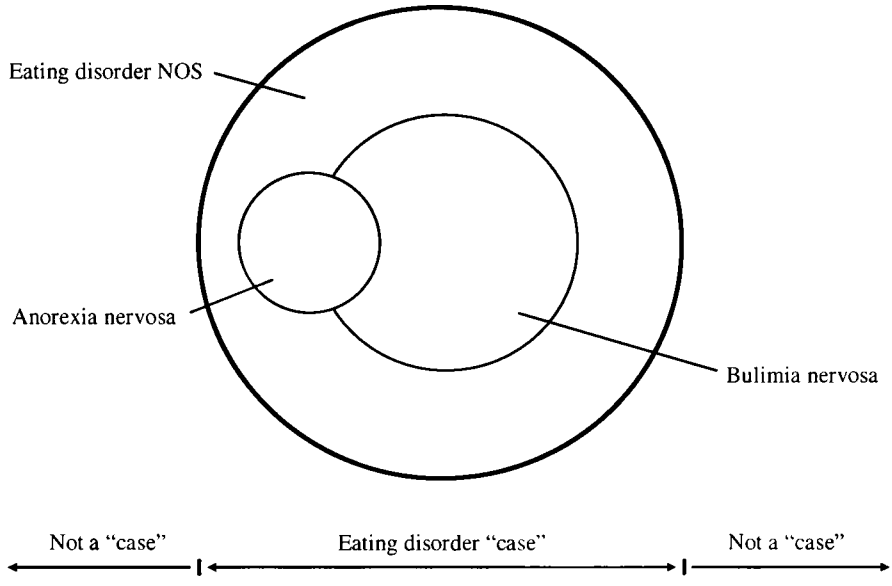
Fairburn and Bohn (the candidate) (2005) proposed the following three solutions to the problems of nosology and neglect that are associated with the diagnosis eating disorder NOS.

4.2.1 Relax the Diagnostic Criteria for Anorexia nervosa and Bulimia nervosa

The first solution was based on the premise that the high prevalence of eating disorder NOS cases is due to the DSM-IV diagnostic criteria for anorexia nervosa and bulimia nervosa being inappropriately strict (see Chapter One, Section 1.3.5). If true, some cases within eating disorder NOS would be better designated as cases of anorexia nervosa or bulimia nervosa.

Figure 1.1 (described in detail in Chapter One) illustrates diagrammatically the relationship between the diagnoses anorexia nervosa, bulimia nervosa and eating disorder NOS. The two overlapping inner circles represent anorexia nervosa and bulimia nervosa respectively. Outside the two inner circles, but within the outer circle, lies eating disorder NOS. Relaxing the diagnostic criteria of anorexia nervosa and bulimia nervosa would involve expanding somewhat the two inner circles.

Figure 1.1 *A schematic representation of the relationship between anorexia nervosa, bulimia nervosa and eating disorder NOS*



Done mindfully, adjusting the diagnostic criteria for anorexia nervosa and bulimia nervosa has much to commend it. As pointed out in Chapter One, many clinicians and researchers have suggested that the DSM-IV criteria need to be adjusted in various ways (Crow et al., 2002; Garfinkel et al., 1995a; Martin et al., 2000; Ramacciotti et al., 2002; Thaw et al., 2001) and in every instance this would have the effect of relaxing the current diagnostic thresholds. Such adjustments seem worth contemplating so long as the two diagnostic concepts are not materially altered. Two main suggestions have been made with respect to anorexia nervosa; the first being that the amenorrhoea criterion be dropped (Cachelin & Maher, 1998; Garfinkel et al., 1996; Watson & Andersen, 2003), and the second being that the "core psychopathology" be redefined to include states in which there is over-evaluation of controlling eating per se without requiring that there also be accompanying concerns about shape and weight (Palmer, 2003; Rieger et al., 2001). Adjusting upward the weight threshold for anorexia nervosa is another option (Garfinkel et al., 1995a; Watson &

Andersen, 2003), although only a marginal change could be accommodated without undermining the fundamental requirement that people with anorexia nervosa be significantly underweight. With regard to bulimia nervosa the main bone of contention concerns the present twice-weekly threshold for the frequency of binge eating: it has been repeatedly argued that a lower minimum frequency would be more appropriate (Garfinkel et al., 1995b; Herzog, Norman, Rigotti, & Pepose, 1986; Wilson & Eldredge, 1991).

Changes of this type represent a fine-tuning of the existing diagnostic criteria rather than any radical change. They involve adding to the two established diagnostic concepts the "subthreshold" cases that exist within eating disorder NOS. This was done by Thaw and colleagues (2001) who systematically modified the DSM-IV criteria of anorexia nervosa and bulimia nervosa in order to determine the impact on the base rates of all three eating disorders. They found that their modifications, in isolation and together, only had a minor impact on the prevalence rate of eating disorder NOS. The clinical relevance of this finding is questionable, however, since Thaw and colleagues used a convenience sample of eating disorder cases, not a clinical one.

In section 4.3 below the impact of systematically adjusting the diagnostic criteria of anorexia nervosa and bulimia nervosa on the relative prevalence of eating disorder NOS, using a representative clinical sample of eating disorder cases, will be examined.

4.2.2 Reclassify Eating Disorder NOS

The second solution proposed by Fairburn and Bohn (2005) was a response to the main shortcoming of the first; namely that it fails to address the fact that many cases within eating disorder NOS are not thought to be subthreshold cases of anorexia nervosa or bulimia nervosa (see Chapter Three, Section 3.3 and section 4.3 below). This solution is an elaboration and extension of the first. Subthreshold cases of anorexia nervosa and bulimia

nervosa would be incorporated within these two diagnoses respectively, as in the first solution, but in addition the remaining cases of eating disorder NOS would be reclassified as belonging to a new category of eating disorder. Clinical experience suggests that the majority of these cases are likely to show the features of anorexia nervosa or bulimia nervosa but combined in different ways, and thus could be described as "mixed" in character, although a minority would fulfil diagnostic criteria for binge eating disorder and might be best separated off. Thus, in summary, this solution involves two steps: the first is to expand anorexia nervosa and bulimia nervosa to embrace the subthreshold cases and the second step is to reallocate the remaining cases to a new diagnostic category, perhaps termed "mixed eating disorder", or to binge eating disorder.

This solution would have the effect of eliminating the concept of eating disorder NOS, at least for the meantime. The diagnosis would re-appear, however, once specific diagnostic criteria for the new diagnostic category were formulated since in practice some "cases" of clinical severity would inevitably be encountered that would fall outside the new boundary, however well it was defined. These cases should be modest in number, however, rendering eating disorder NOS a small residual category, as NOS categories are intended to be.

4.2.3 The "Transdiagnostic" Solution

The third solution suggested by Fairburn and Bohn (2005) was the most radical. It was to bring eating disorder NOS into the limelight by creating a single unitary diagnostic category "eating disorder" embracing anorexia nervosa, bulimia nervosa and eating disorder NOS without any subdivisions. The main argument for proposing a "transdiagnostic" solution of this type is that the current emphasis on subdividing the eating disorders (into anorexia nervosa and bulimia nervosa, each with their two subtypes, eating disorder NOS

and possibly binge eating disorder) detracts attention from the most striking characteristic of the eating disorders; namely, that far more unites the various forms of eating disorder than separates them (Fairburn & Harrison, 2003; Waller, 1993; Walsh & Garner, 1997). Thus, rather than focusing on differences between the eating disorders, there is a case for highlighting the many features that are shared by them and are largely peculiar to them. These include extreme dietary restraint and restriction, binge eating, self-induced vomiting and the misuse of laxatives, driven exercising, body checking and avoidance, and the over-evaluation of control over eating, shape and weight. These cross-diagnostic similarities become even more obvious if a longitudinal perspective is taken since patients do not adhere to their DSM-IV diagnosis over time; rather, they move between them (Fairburn and Harrison, 2003; Herzog et al, 1993; Milos et al, 2005).

Fairburn and Bohn (2005) argued that in the longer term the transdiagnostic solution might be the most valuable one, but that the second solution might have most to recommend it in the short term, given that the first solution ignores the fact that many of the cases within eating disorder NOS seem to be of the mixed variety (see section 4.3 in this chapter). The second proposal has therefore also been termed the "Interim Proposal" (see Section 4.4 in this chapter).

4.2.4 Summary

- Three solutions to the problems of nosology and neglect that are associated with the diagnosis eating disorder NOS were proposed by Fairburn and Bohn (2005).
- The first solution was based on the premise that the high prevalence of eating disorder NOS cases is due to the DSM-IV diagnostic criteria for anorexia nervosa and bulimia nervosa being inappropriately strict. In involved systematically adjusting the diagnostic criteria for anorexia nervosa and bulimia nervosa in order

to embrace their subthreshold cases currently within eating disorder NOS. Fairburn and Bohn (2005) suggest that the adjustments should be such that the core diagnostic concepts of anorexia nervosa and bulimia nervosa are not altered.

- The second solution was an elaboration and extension of the first. It involved two steps: the first was to adjust the diagnostic criteria of anorexia nervosa and bulimia nervosa as described in solution one, and the second was to reallocate the remaining cases (of which there were thought to be many) to a new diagnostic category, perhaps termed "mixed eating disorder" or to binge eating disorder.
- The third solution was to create a single unitary diagnostic category "eating disorder" embracing anorexia nervosa, bulimia nervosa and eating disorder NOS without any subdivisions. This transdiagnostic solution was based on the premise that the three eating disorders are far more similar than dissimilar.

4.3 The impact of adjusting the DSM-IV criteria for Anorexia nervosa and Bulimia nervosa on the relative prevalence of Eating Disorder NOS

4.3.1 Introduction and aim

As outlined in Chapter One many researchers and clinicians have stated that the diagnostic criteria for anorexia nervosa and bulimia nervosa are too narrow and therefore do not match clinical reality. Suggestions have been made as to how best to adjust the diagnostic criteria for the two specified diagnoses so that the number of cases classified as eating disorder NOS is reduced (see Chapter One, Section 1.3.5). To date only one study has systematically lowered the diagnostic thresholds of anorexia nervosa and bulimia nervosa and found that the impact on the prevalence of eating disorder NOS was minor

(Thaw et al., 2001). This study, however, did not use a clinical sample of eating disorder cases.

The aim of the current study was to evaluate the impact of adjusting the DSM-IV criteria for anorexia nervosa and bulimia nervosa on the relative prevalence of eating disorder NOS, using a representative sample of adult outpatients with an eating disorder of clinical severity. The study was a test of the clinical usefulness of the first of three solutions to the problems of eating disorder NOS, described in section 4.2 of this chapter. It was supposed to answer the question whether the anomalously high relative prevalence of eating disorder NOS in adult outpatient settings might be due in part to the inclusion within it of cases closely resembling anorexia nervosa and bulimia nervosa, cases that might be better reclassified as such.

4.3.2 Methods

Design

To achieve the above mentioned aim it was necessary to assess a representative clinical sample of eating disorder cases using an instrument that generates eating disorder diagnoses and the type of data that allowed the diagnostic criteria for anorexia nervosa and bulimia nervosa to be adjusted in a systematic way. The Eating Disorder Examination (EDE) interview (Fairburn & Cooper, 1993) was chosen for this purpose. The diagnostic criteria for anorexia nervosa and bulimia nervosa were systematically relaxed along lines advocated in the literature (see Chapter One, Section 1.3.5), while still preserving the core clinical concepts (e.g., that patients with anorexia nervosa be actively maintaining an objectively low weight, and that patients with bulimia nervosa experience repeated episodes of binge eating). After each adjustment the relative prevalence rates for eating disorder NOS, anorexia nervosa and bulimia nervosa were recalculated.

Participants

The sample comprised the 170 patients with a clinical eating disorder described in Chapter Three. The patients were consecutive referrals to two outpatient eating disorder clinics, one in Oxfordshire and the other in Leicestershire. They had been referred for the treatment of their eating disorder. Patients were assessed by one of three senior clinicians (Dr Zafra Cooper or Professor Christopher Fairburn in Oxford; Professor Robert Palmer in Leicester) who determined whether the patient had an eating disorder of clinical severity. They were subsequently assessed by one of six trained research assistants (Marianne O'Connor, Caroline Plumb and Shani Langdon in Oxford; Jackie Wales, Claire Nollett and Elizabeth Benson in Leicester) using various measures including the EDE. The nature of the sample (i.e., demographic features, clinical characteristics) is described in detail in Chapter Three, Section 3.2.

Measures

Operational eating disorder diagnoses were created using the Eating Disorder Examination (EDE; Fairburn & Cooper, 1993). The EDE has been discussed in detail in Chapter Three, Section 3.2.2. The DSM-IV criteria for anorexia nervosa and bulimia nervosa and their EDE-based operational definitions are shown in Tables 3.1 and 3.2 respectively (see Chapter Three, p.67-68). As specified by DSM-IV, those eating disorders that did not meet these two operational definitions were classed as cases of eating disorder NOS. In addition, diagnoses of "binge eating disorder" were made based on the research criteria specified in DSM-IV (see Table 1.4 in Chapter One, p.17). The "Binge eating disorder module" of the EDE (see Appendix 3.1) was used to assess whether the research criteria for binge eating disorder were fulfilled. Binge eating disorder is a provisional new

eating disorder diagnosis. Technically, it is an example of the type of clinical problem subsumed under the rubric of eating disorder NOS (see Chapter One, Section 1.3.4).

Procedure

Three diagnostic criteria for anorexia nervosa and one diagnostic criterion for bulimia nervosa were systematically modified one-by-one and then in combination. The diagnostic criteria for anorexia nervosa were altered in the following way:

Change 1 the amenorrhoea criterion (criterion D) was omitted;

Change 2 the weight threshold (criterion A) was adjusted upwards from a BMI of 17.5 to a BMI of lower or equal 18.0

Change 3 the weight threshold (criterion A) was raised still further to a BMI of lower or equal 18.5

Change 4 an alternative criterion to criteria B and C was added, criterion E: this was the over-evaluation of controlling eating per se. It was defined by a rating of 4 or higher on the EDE item "Importance of controlling eating". Patients who fulfilled this criterion, but not criteria B and C, were included in the diagnosis of anorexia nervosa. (Patients who fulfilled criteria B and C but not E retained their anorexia nervosa diagnosis.)

One change was made to the diagnostic criteria for bulimia nervosa. The frequency threshold for objective bulimic episodes and compensatory behaviour (criterion C) was lowered from an average of at least twice per week to an average of at least once per week. The time window of three months was retained. Thus, patients were eligible for the (expanded) diagnosis of bulimia nervosa if they reported both, 12 or more objective bulimic episodes over the past three months and four or more in the past month, together with 12 or more episodes of purging (self-induced vomiting, laxative misuse or diuretic misuse) over the past three months, and four or more episodes of purging in the past month.

It should be noted that the DSM-IV 'trumping rule' for anorexia nervosa was retained; that is, the diagnosis of anorexia nervosa took precedence over that for bulimia nervosa. This meant that, if the criteria for both diagnoses were fulfilled, a diagnosis of anorexia nervosa was made.

Data analysis

Patients' EDE data were entered into SPSS version 14.0 and the diagnostic criteria were modified as described accordingly. The relative prevalence rates of each eating disorder diagnosis were re-calculated after each modification.

4.3.3 Results

The diagnostic distribution of the sample can be seen in Figure 3.1, taken from Chapter Three. Of the 170 patients with a clinical eating disorder, eight (4.7%) fulfilled DSM-IV criteria for anorexia nervosa, 60 (35.3%) met diagnostic criteria for bulimia nervosa and 102 (60.0%) were classified as eating disorder NOS. Seven of the 102 eating disorder NOS cases (6.9%) fulfilled the (provisional) DSM-IV criteria for binge eating disorder.

Figure 3.1 *Relative prevalence rates of the three eating disorders in a sample of 170 outpatients*

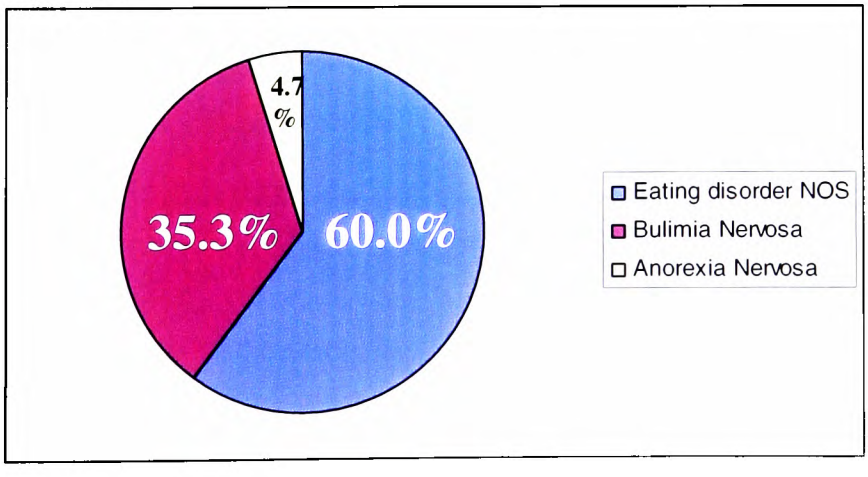


Table 4.1 shows the impact of each of the five changes on the relative prevalence of eating disorder NOS, anorexia nervosa and bulimia nervosa. They are described for each change in turn.

Table 4.1 *Impact on the relative prevalence of the three DSM-IV eating disorders of relaxing the diagnostic criteria for Anorexia Nervosa and Bulimia Nervosa*

	Anorexia Nervosa N (%)	Bulimia Nervosa N (%)	Eating Disorder NOS N (%)
1. Current DSM-IV diagnostic criteria	8 (4.7)	60 (35.3)	102 (60.0)
2. Adjustments to the anorexia nervosa criteria:			
a. Removal of the amenorrhea criterion	12 (7.1)	58 (34.1)	100 (58.8)
b. Raising of the BMI threshold to ≤18.0 (kg/m^2)	8 (4.7)	60 (35.3)	102 (60.0)
c. Raising of the BMI threshold to ≤18.5 (kg/m^2)	12 (7.1)	58 (34.1)	100 (58.8)
d. Addition of those who overvalue control over eating *per se*	8 (4.7)	60 (35.3)	102 (60.0)
e. Adjustments a, c and d combined	21 (12.4)	56 (32.9)	93 (54.7)
3. Adjustments to the bulimia nervosa criteria:			
Reduction of the minimum average frequency of binge eating and purging to at least once per week	8 (4.7)	68 (40.0)	94 (55.3)
4. Adjustments to the anorexia nervosa and bulimia nervosa criteria combined	21 (12.4)	64 (37.6)	85 (50.0)

Adjusting the diagnostic criteria for anorexia nervosa

Modifications in isolation:

Change 1: Removing the amenorrhoea criterion from the diagnostic criteria for anorexia nervosa only had a minor impact on the prevalence rate of eating disorder NOS. The prevalence rate of eating disorder NOS was reduced by 1.2%.

Change 2: Raising the weight threshold for anorexia from a body mass index of lower or equal 17.5 to lower or equal 18.0 did not affect the relative prevalence of eating disorder NOS. Nor did it have any impact on the prevalence rates of the other two eating disorders.

Change 3: Raising the weight threshold for anorexia nervosa even further to a body mass index of lower or equal 18.5 had hardly any impact at all on the prevalence of eating disorder NOS. The prevalence rate of eating disorder NOS was reduced by 1.2%.

Change 4: Also including patients who were primarily concerned with controlling their eating (i.e., they did not over-evaluate shape and/or weight and were not afraid of weight gain) did not have any impact at all on any of the three prevalence rates.

Modifications in combination

Removing the amenorrhoea criterion, adjusting the weight threshold to a body mass index of 18.5 and including people who were overly concerned with controlling eating per se into the diagnosis of anorexia nervosa only had a marginal impact on the prevalence rates of the three disorders. More than 50% of the sample retained their eating disorder NOS diagnosis. Only nine former eating disorder NOS cases were now classified as anorexia nervosa. The relative prevalence rate of eating disorder NOS was thereby reduced by 5.3%. In addition, four cases of bulimia nervosa were now classified as anorexia nervosa.

Adjusting the diagnostic criteria for bulimia nervosa

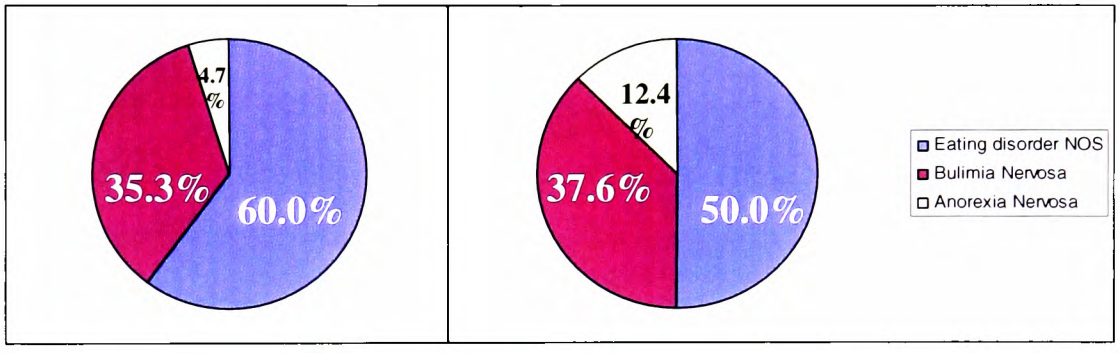
Change 1: Lowering the threshold for objective bulimic episodes and compensatory behaviour to an average of once per week had an equally minor effect on the prevalence rate of eating disorder NOS: it was reduced by only 4.7%. Eight former eating disorder NOS cases now fulfilled the diagnostic criteria of bulimia nervosa.

Combining the changes to the diagnostic criteria for anorexia nervosa and bulimia nervosa:

Combining all the changes to diagnostic criteria of anorexia nervosa and bulimia nervosa with the weight threshold for anorexia nervosa being set at a body mass index of lower or equal 18.5 reduced the prevalence rate of eating disorder NOS by only 10%. The

relative prevalence of bulimia nervosa was marginally increased (by 2.3%), as was that of anorexia nervosa (by 7.7%). Yet, half the sample still retained the diagnosis eating disorder NOS (nine of which were now classified as binge eating disorder). Figure 4.1 shows the prevalence rates of the three eating disorders before and after all five modifications to the diagnostic criteria of anorexia nervosa and bulimia nervosa were made. It can be seen that the change was minor.

Figure 4.1 *Relative prevalence rates of the three eating disorders in a sample of 170 outpatients before (left) and after (right) adjusting the diagnostic criteria of anorexia nervosa and bulimia nervosa in combination*



4.3.4 Summary of findings

- In a representative sample of 170 outpatients with an eating disorder of clinical severity the following diagnostic distribution was found: eight patients (4.7%) fulfilled DSM-IV criteria for anorexia nervosa, 60 (35.3%) met diagnostic criteria for bulimia nervosa and 102 (60.0%) were classified as eating disorder NOS.
- The DSM-IV criteria of anorexia nervosa and bulimia nervosa were adjusted in a systematic way and the impact of these adjustments on the relative prevalence of eating disorder NOS was examined.
- The following changes were made to the diagnostic criteria for anorexia nervosa: the amenorrhoea criterion was dropped, the weight threshold was raised from a body mass index of 17.5 to 18.0 and in a second step to 18.5, and the core

psychopathology was changed such as to include people who are primarily concerned with controlling eating per se.

- With regard to bulimia nervosa, the only change made was the lowering of the minimum frequency of objective bulimic episodes and compensatory behaviours from twice per week to once per week.
- It emerged that none of the adjustments either in isolation or in combination had much impact on the relative prevalence of eating disorder NOS which remained at fifty percent or more.
- Nor was the presence of binge eating disorder responsible for the large number of eating disorder NOS cases for just seven patients met its diagnostic criteria. [This figure rose to nine if the minimum average frequency of binge eating was reduced to once per week.] The low prevalence of binge eating disorder was not unexpected as three of the five patient samples referred to above (Section 4.1) reported prevalence figures for binge eating disorder and in all three instances the prevalence was less than ten percent (Martin et al., 2000; Nollett & Button, 2005; Ricca et al., 2001).

4.4 Evaluating the Clinical Utility of Fairburn and Bohn's (2005) "Interim Proposal"

4.4.1 Introduction

In a recent article in the *American Journal of Psychiatry* First and colleagues (First et al., 2004) discuss the concept of "clinical utility" and its relevance to diagnosis. They defined "clinical utility" as "the extent to which DSM (the American Psychiatric Association's Diagnostic and Statistical Manual for Mental Disorders) assists clinical decision makers in fulfilling the various clinical functions of a psychiatric classification system". They suggest that these functions include assisting clinicians with 1)

conceptualising diagnostic entities, 2) communicating clinical information to practitioners and patients and their families, 3) using diagnostic categories and criteria sets in clinical practice, 4) choosing effective interventions, and 5) predicting future clinical management needs. First and colleagues particularly stress the importance of determining whether changes to diagnostic criteria improve their clinical utility. Doing so, they argue, would clarify whether the advantages of changing the diagnostic criteria outweigh potential negative consequences. They state that in the revision process for the fourth edition of DSM, efforts were made to improve the validity of diagnostic criteria sets, using comprehensive literature reviews, data reanalyses and field trials, but no effort was put into empirically examining whether any changes proposed actually improved clinical utility. This was regrettable given the statement in the introduction to DSM-IV (APA, 1994), that its "highest priority has been to provide a helpful guide to clinical practice" (APA, 1994, p.xv).

First and colleagues (2004) suggest that, based on their definition of clinical utility, the merits of any proposed change to DSM could be evaluated by considering 1) its impact on the use of DSM, 2) whether it enhances clinical decision making, and 3) whether it improves clinical outcomes. One way of determining whether a proposed change has an impact on the use of DSM, is, according to First and colleagues, to assess how acceptable it is to users. This can be done by surveying users' opinions on a proposed change, especially with regard to the validity of the change and its potential ease of use in clinical practice.

The aim of this part of the present research was to examine, using a survey method, whether Fairburn and Bohn's (2005) proposed changes to the diagnostic criteria of anorexia nervosa, bulimia nervosa and eating disorder NOS were acceptable to users of these diagnostic criteria, thereby evaluating one aspect of their clinical utility. The specific change

studied was the second of Fairburn and Bohn's (2005) three solutions to the problem of eating disorder NOS, the so-called "Interim Proposal". In summary it involved two steps:

1. Expanding the diagnostic criteria of anorexia nervosa and bulimia in order to embrace "subthreshold" cases of AN and BN currently located within eating disorder NOS, thereby removing them from the diagnostic category; and
2. Reclassifying the remaining cases of eating disorder NOS, either to a new diagnostic category, perhaps termed "mixed eating disorder" or to binge eating disorder.

4.4.2 Aim

The aim of this study was to evaluate, by means of a survey, the opinion of eating disorder experts on:

1. the DSM-IV diagnosis eating disorder NOS, and the associated problems of nosology and neglect as suggested by Fairburn and Bohn (2005); and
2. the second of three solutions to these problems (the 'Interim proposal') proposed by Fairburn and Bohn (2005), in terms of its likely validity, acceptability, ease of use and potential.

4.4.3 Methods

Design

The Fairburn and Bohn (2005) article, together with a 17-item questionnaire, was sent via e-mail to the 190 people who were registered to attend the 2004 annual meeting of the Eating Disorder Research Society in Amsterdam. The two documents were sent one month before the meeting was due to take place. Recipients were asked to read the article and then complete the questionnaire which was designed to assess their views on eating disorder NOS and Fairburn and Bohn's (2005) proposed solution. They were also invited to make any further comments if they wished.

Participants

Those due to attend the annual meeting of the Eating Disorder Research Society were regarded as a suitable study sample since the meeting is only open to experts on eating disorders and their guests. The participants were therefore likely to be knowledgeable about eating disorders and the problems associated with the DSM-IV diagnostic scheme.

Distribution of the Questionnaire

The questionnaire was sent to the participants via email by the organising committee of the meeting. Recipients were told that they could remain anonymous if they wished by returning the questionnaire indirectly via the organising committee (who would pass it on anonymously) rather than returning it directly to the candidate.

Content of the Questionnaire

The questionnaire is shown in Table 4.2. It was composed jointly by the candidate and Professor Fairburn. It consisted of 17 questions and a 4-point Likert scoring scale for each item, assessing the participants' level of agreement. The four Likert responses were as follows: "Do not agree at all", "Agree somewhat", "Agree to a large extent", and "Completely agree" Participants were asked to leave a question blank if they could not answer it. The 17 items comprised:

1. Five questions regarding the DSM-IV diagnosis eating disorder NOS, cases of eating disorder NOS, and the problems of nosology and neglect that Fairburn and Bohn (2005) stated were associated with this diagnosis
2. Twelve questions regarding Fairburn and Bohn's 'Interim proposal' (2005). These consisted of
    i. One question concerning the respondents' overall view on the proposed change

ii. Six questions about whether the proposed change addressed the problems associated with eating disorder NOS

iii. Two questions about the diagnostic validity of the proposed change

iv. Three questions regarding the likely ease of use of the proposed change.

The cover page contained brief instructions on how to complete the questionnaire and a synopsis of the proposed diagnostic change. The last page contained a section for comments and three questions about the participants' level of clinical experience working with patients with eating disorder and their professional background. The entire document, as sent to potential participants, is shown in Appendix 4.1.

Table 4.2 *Self-report Questionnaire sent to the registrants of the 2004 annual meeting of the Eating Disorder Research Society in Amsterdam*

INSTRUCTIONS
Please place an 'X' in the column which best describes your response to the following questions.

Thank you.

	If you cannot answer any question please leave it blank.	Do not agree at all	Agree somewhat	Agree to a large extent	Completely agree
1	Do you agree that it is a problem that EDNOS is a "residual" diagnostic category?				
2	Do you agree (on the basis of your clinical experience) that EDNOS is common?				
3	Do you agree that EDNOS is neglected in terms of clinical guidance and research?				
4	Do you agree (on the basis of your clinical experience) that many cases of EDNOS are of clinical severity equivalent to bulimia nervosa?				

	If you cannot answer any question please leave it blank.	Do not agree at all	Agree somewhat	Agree to a large extent	Completely agree
5	Do you agree (on the basis of your clinical experience) that there is heterogeneity within EDNOS?				
6	Do you think that the Interim Proposal of Fairburn and Bohn has merit to it (i.e., Do you like it?)?				
7	Do you think that the Interim Proposal addresses some (or all) of the problems associated with EDNOS?				
8	Do you think that Step 1 (i.e., expanding somewhat the diagnostic criteria for AN and BN) would be sufficient on its own to address the problems listed above?				
9	Do you think that the addition of Step 2 is needed to address the problems listed above (i.e., reclassifying the remaining cases within EDNOS as "mixed eating disorder" or binge eating disorder)?				
10	Do you think the Interim Proposal (with the creation of four eating disorder diagnoses; AN, BN, mixed eating disorder and BED) matches clinical reality as you see it in your practice?				
11	Do you think the Interim Proposal would make the diagnostic system for eating disorders too complex?				
12	Do you think the Interim Proposal would be usable in practice?				
13	Do you think it would be helpful for some of your EDNOS patients to be given the new diagnosis "mixed eating disorder"?				
14	Do you think the full Interim Proposal would draw attention to EDNOS patients with regard to defining their clinical characteristics?				
15	Do you think the full Interim Proposal would draw attention to EDNOS patients with regard to defining their clinical needs?				
16	Do you think the full Interim Proposal would draw attention to EDNOS patients with regard to encouraging research on their treatment?				
17	Do you think the Interim Proposal is an improvement over the existing DSM-IV diagnostic scheme for eating disorders?				

4.4.4 Results

Respondents

Sixty-eight of the 190 meeting registrants (35.8%) returned the questionnaire. Amongst the 68 respondents, there were 26 psychiatrists, 22 clinical psychologists, nine physicians, five doctoral students in clinical psychology, five university academics and one nurse. Sixty-three (92.6%) were currently working clinically with patients with eating disorders and they had done so for, on average, 14.9 years (SD 10.4 years; range: 1-40 years).

Respondents' views regarding the problems associated with eating disorder NOS

The respondents' answers to the five questions about their views on the diagnosis eating disorder NOS and patients with this diagnosis are shown in Table 4.3 below. As can be seen, the majority (between 73.6% and 92.7%) agreed with Fairburn and Bohn (2005)[6] with respect to the problems associated with eating disorder NOS. Specifically, over ninety percent (91.2%) thought that it was a problem that eating disorder NOS is a "residual" diagnostic category; 92.6% agreed that eating disorder NOS is common; 92.7% agreed that eating disorder NOS is neglected in terms of clinical guidance and research; and 73.6% agreed that there is heterogeneity within eating disorder NOS.

Respondents' views regarding the 'Interim proposal' (Fairburn and Bohn, 2005)

The respondents' answers to the questions about Fairburn and Bohn's (2005) proposed changes to the diagnostic system and their potential for addressing the problems of nosology and neglect are shown in Table 4.4 below.

[6] In this summary of the findings the term "agree" is used for responses that belong to either of the categories "Agree to a large extent" or "Completely agree"

Table 4.3 *Respondents' answers to questions regarding Eating Disorder NOS*

Question	N	Do not agree at all N (%)	Agree somewhat N (%)	Agree to a large extent N (%)	Completely agree N (%)
Do you agree that it is a problem that EDNOS[1] is a "residual" diagnostic category?	68	1 (1.5)	5 (7.4)	14 (20.6)	48 (70.6)
Do you agree (on the basis of your clinical experience) that EDNOS is common?	66		3 (4.4)	9 (13.2)	54 (79.4)
Do you agree that EDNOS is neglected in terms of clinical guidance and research?	68		5 (7.4)	18 (26.5)	45 (66.2)
Do you agree (on the basis of your clinical experience) that many cases of EDNOS are of clinical severity equivalent to BN[2]?	64	2 (2.9)	12 (17.6)	15 (22.1)	35 (51.5)
Do you agree (on the basis of your clinical experience) that there is heterogeneity within EDNOS?	65		2 (2.9)	12 (17.6)	51 (75.0)

[1] EDNOS = Eating disorder NOS
[2] BN = Bulimia Nervosa

Table 4.4 *Respondents' answers to questions regarding Fairburn and Bohn's 'Interim proposal' and its potential for addressing the problems associated with Eating Disorder NOS*

Question	N	Do not agree at all N (%)	Agree somewhat N (%)	Agree to a large extent N (%)	Completely agree N (%)
Do you think that the Interim Proposal of Fairburn and Bohn has merit to it (i.e., Do you like it?)?	68	3 (4.4)	12 (17.6)	34 (50.0)	19 (27.9)
Do you think that the Interim Proposal addresses some (or all) of the problems associated with EDNOS?	67	3 (4.4)	23 (33.8)	27 (39.7)	14 (20.6)

(Table 4.4 continued)

Question	N	Do not agree at all N (%)	Agree somewhat N (%)	Agree to a large extent N (%)	Completely agree N (%)
Do you think that Step 1 (i.e., expanding somewhat the diagnostic criteria for AN[1] and BN) would be sufficient on its own to address the problems listed above?	68	35 (51.5)	23 (33.8)	8 (11.8)	2 (2.9)
Do you think that the addition of Step 2 is needed to address the problems listed above (i.e., reclassifying the remaining cases within EDNOS as "mixed eating disorder" or binge eating disorder)?	67	2 (2.9)	16 (23.5)	35 (51.5)	14 (20.6)
Do you think the full Interim Proposal would draw attention to EDNOS patients with regard to defining their clinical characteristics?	68		16 (23.5)	37 (54.4)	15 (22.1)
Do you think the full Interim Proposal would draw attention to EDNOS patients with regard to defining their clinical needs?	68	2 (2.9)	26 (38.2)	32 (47.1)	8 (11.8)
Do you think the full Interim Proposal would draw attention to EDNOS patients with regard to encouraging research on their treatment?	68		13 (19.1)	30 (44.1)	25 (36.8)

[1] AN = Anorexia Nervosa

More than two-thirds (77.9%) thought that the proposed changes had merit to them, and 60.3% thought that the proposed changes would address some or all of the problems associated with the diagnosis. Few (14.7%) thought that Step 1 of the proposed changes (relaxing somewhat the diagnostic criteria for anorexia nervosa and bulimia nervosa) would be sufficient on its own to address the problems whereas three-quarters (72.1%) thought that the addition of Step 2 (reclassifying the remaining cases as "mixed eating disorder" or BED) would be likely to do so. Between 58.9% and 80.9% of the respondents agreed that the full Interim proposal would draw attention to eating disorder NOS patients with regard to

defining their clinical characteristics and their clinical needs, and encouraging research on their treatment.

Respondents' view regarding the validity and likely ease of use of the 'Interim proposal'

Respondents' answers to the five questions about the validity and the likely ease of use of the proposed changes are shown in Table 4.5.

Table 4.5 *Respondents' answers to questions regarding the validity and the likely ease of use of the 'Interim proposal' (Fairburn and Bohn, 2005)*

Question	N	Do not agree at all N (%)	Agree somewhat N (%)	Agree to a large extent N (%)	Completely agree N (%)
Do you think the Interim Proposal (with the creation of four eating disorder diagnoses; AN, BN, mixed eating disorder and BED[1]) matches clinical reality as you see it in your practice?	65	2 (2.9)	16 (23.5)	38 (55.9)	9 (13.2)
Do you think the Interim Proposal would make the diagnostic system for eating disorders too complex?	68	56 (82.4)	7 (10.3)	4 (5.9)	1 (1.5)
Do you think the Interim Proposal would be usable in practice?	66	2 (2.9)	11 (16.2)	40 (58.8)	13 (19.1)
Do you think it would be helpful for some of your EDNOS patients to be given the new diagnosis "mixed eating disorder"?	62	9 (13.2)	22 (32.4)	18 (26.5)	13 (19.1)
Do you think the Interim Proposal is an improvement over the existing DSM-IV diagnostic scheme for eating disorders?	67	1 (1.5)	10 (14.7)	30 (44.1)	26 (38.2)

[1] BED = Binge Eating Disorder

Over two-thirds (69.1%) agreed that the proposed changes matched clinical reality as they see it in their clinical practice. Over eighty percent (82.4%) did not think that the proposed changes would make the diagnostic system for eating disorders too complex and

three-quarters (77.9%) thought that they would be usable in practice. On the other hand, less than half (45.6%) thought that it would be helpful for some of their eating disorder NOS patients to be given the diagnosis "mixed eating disorder". Over eighty percent (82.3%) thought that the proposed changes would be an improvement over the existing DSM-IV diagnostic scheme.

Additional comments from respondents

Twenty-four of the 68 respondents (35.3%) provided additional comments on their questionnaires. These are shown in Table 4.6. It can be seen that the comments contained a broad range of opinions regarding Fairburn and Bohn's (2005) 'Interim proposal'. Fourteen of the 24 comments (58.3%) were generally positive, four of which welcomed the 'Transdiagnostic solution' (see Section 4.2.3 of this chapter) as the ultimate step forward. Eight of the 24 comments (33.3%) expressed doubts about the clinical usefulness of the Interim proposal, five of which strongly favoured the 'Transdiagnostic solution'. Two respondents (8.3%) provided general comments about classification systems.

Table 4.6 *Additional comments provided by the respondents*

Comments generally in favour of the 'Interim proposal' (Fairburn and Bohn, 2005):

1. "Great paper – will be important in stimulating the debate and hopefully prompting more action in terms of moving forward from the current hopelessly inadequate situation. In some senses does not feel radical enough, but probably a very sensible first step proposal. Could be really important in getting people to think more about the simple but obvious fact that clinical need is not always clearly indicated by diagnosis. I think there remains value in approaching the whole issue of what works for whom from a range of different perspectives (i.e. not just presenting clinical features of ED and diagnostic categories), but that is not incompatible with what you propose, and a healthy degree of realism is prudent if your ideas are to have maximum impact. Thanks for your work in taking this forward.
2. "Thank you; this is just what we need as a first step. In the future I would be in favour of considering more complexity instead of one "mixed" category. For example the non-bingeing purgers who do not meet criteria for AN may (or may not) form a justifiably separate category. We could develop criteria for deciding whether any of the specific subgroups should qualify as a diagnostic entity (e.g., one criterion might be that the pattern occurs in a sufficient number of individuals)."

3. "Problems in the diagnostic criteria of eating disorders are among the most significant challenges we face as a field. This manuscript is an outstanding summary of the problems with the classification system and provides several useful suggestions to changing the current criteria. The proposed modifications to the DSM-IV criteria are appropriate and potentially beneficial in terms of both research and clinical work.
   In spite of the appropriateness of these diagnostic changes, I believe that the proposed changes would pose some problems with differential diagnosis. For example, would individuals who purge but report SBEs and no OBEs be counted as mixed ED or bulimia nervosa? What about patients with anorexia nervosa who are partially weight restored? As discussed in the end of the paper, the importance of determining a threshold level of symptoms for "caseness" in the new mixed criteria will also be critical. Nonetheless, I fully support the authors' suggestions and hope that the next version of the DSM will incorporate them. Although I understand the advantages of a transdiagnostic system, this approach is highly inconsistent with the categorical nature of the DSM and not likely to be adopted in the US in the near future."
4. "The criteria for the "mixed" group would have to be clear and readily assessed to be of use. Given that the provisional identification of BED in the DSM-IV has led to treatment & diagnostic research on this subgroup, it seems likely that reclassifying EDNOS would be very helpful in generating research/writing on this ignored group."
5. "I think the interim proposal has merit but in judging any diagnostic scheme I would look to try to identify some markers of validity. In my clinical work the status of EDNOS has little bearing, but in my research life, a huge impact."
6. "Well articulated and argued. Essentially it seems logical, in that other than BED, the other cases are indeed "mixed". Not sure what it gets us, other than getting past the problem of the most frequent pattern – i.e., 'mixed', being the most common. But I understand the point – doing so at least gets this substantial proportion of cases onto the playing field as far as research goes."
7. "Main problem is for researchers, most of us already clinically classify folks these ways- but a good idea to formalize this. Considerable information already exists on issues surrounding the severity of individuals with partial syndrome AN and BN. Think BED continues to warrant investigation as a separate diagnosis."
8. "My comments pertain to the mixed eating disorder group, as this will still be quite a large proportion of clients seen (about a third). I think we need a clear strategy for rapidly moving beyond the "interim state" with this group, and (a) using latent class analyses to examine any logical groupings (e.g., do we really need a separate BED group), (b) examine psychopathology compared to ED groups, (c) and decide on what outer limits define this group. I guess I am suggesting that an Interim proposal is a good idea, but there needs to be a clearly articulated strategy for moving beyond at the same time, so there is a cohesive direction of research laid down to keep things moving."
9. "I read the paper and agree that something has to be done. I even like the third idea of just a category called Eating Disorder."
10. "The attempt to define new criteria for EDNOS is very important."

Comments in favour of the 'Interim proposal', but preferring the 'Transdiagnostic solution':

11. "Although I applaud the Interim Proposal as a much needed step, I am completely in favor of the "transdiagnostic" solution, based on my knowledge about the course of eating disorders."
12. "I think that the full Interim Proposal could be useful to increase research on the neglected EDNOS category. I am doubtful that it will add some advantage to the clinical practice. I think that the "transdiagnostic solution" is potentially more clinically powerful and much easier to implement in the real world than a system that includes 4 diagnostic categories."

13. "I strongly support the transdiagnostic unitary category. I think the Interim Proposal is a very good compromise and a big improvement compared to the current system. I hope it will not postpone even more the adoption of the transdiagnostic perspective though. I have some doubts about the "mixed eating disorders" diagnosis."
14. "I think that this Interim Proposal is the first step to deal with the EDNOS patients but the final solution will be the Transdiagnostic."

Comments doubting the clinical usefulness of the 'Interim proposal' and/or favouring the 'Transdiagnostic solution':

15. "Personal viewpoint: My concern is that this is another half-way house (although a sensible one, if one is to have a half-way house). It is the notion of diagnoses in the eating disorders that is flawed: not their details or their application. To formulate and treat, it is more useful to focus on the traits of restriction and bulimia than it is to try to create artificial disjunctions. Diagnosis remains a shorthand for the use of clinicians, and a lazy one at that – not a case of 'cutting nature at its joints'. The transdiagnostic approach is the sensible one, and always has been – having diagnoses within that framework is a category error of the type that Gilbert Ryle would have railed at."
16. "My choice is the transdiagnostic solution. It may seem radical, but it is analogue to what happened in depression research 20 years ago. My proposal "Major Eating Disorder" and AN, BN and BED as 'subtypes' "
17. "One category of eating disorder and dimensional assessment of symptoms is probably the ultimate solution."
18. "I am concerned that the interim proposal is a band-aid solution because it does not do much beyond changing the name of the EDNOS category, which will remain an extremely heterogeneous group. Also, the addition of a mixed category might also imply that AN and BN themselves are non-mixed. In actuality, binge/purge type AN could also be viewed as "mixed" eating disorder in that it contains symptoms from both AN and BN. Finally, use of the term "mixed eating disorder" implies that it is a completely separate entity from AN and BN, and, as is discussed in the paper, the current DSM-IV distinction between AN/BN and EDNOS is extremely arbitrary.
    I think that more data needs to be collected on the potential differences in severity of eating pathology and general psychopathology between AN/BN and EDNOS. Having read several of these studies, I have observed that although the differences found are often not statistically significant, effect sizes sometimes suggest that differences — albeit statistically non-significant — may exist. This makes me wonder if some of these studies are underpowered, and underscores a problem with current statistical techniques: one cannot confirm the null hypothesis.
    I very much prefer the transdiagnostic approach proposed at the end of the paper. I agree that the current diagnostic categories are a historical accident and that perhaps we need to start from scratch. As is pointed out in the paper, eating disorders share many important similarities and multiple studies show substantial crossover among diagnoses over time. I think this is lost in the current DSM-IV diagnostic scheme. I have also been very impressed with the treatment outcome results your group has presented on your transdiagnostic approach to treatment, which speaks to the potential clinical utility of a transdiagnostic classification system."
19. "I would prefer a "transdiagnostic solution" and an integration of 'stages' according to impairment/severity, duration of illness and change of diagnostic criteria over time."
20. "I think that shifts in diagnostic criteria should be gradual in order to evaluate their impact on research and treatment. Shifts should have compelling reasons (e.g. treatment response differs for diagnostic subgroups; prognosis varies). Just now although diagnoses may vary over time for a significant subset, effective treatments for AN and BN appear to differ."

21. "In Germany, the problem is not so striking, because all have to make diagnosis in ICD-10. Here, the criteria are not so strict, and there are the categories atypical anorexia nervosa and atypical bulimia nervosa. This reduces the number of "other eating disorders", which might be similar to EDNOS. Perhaps it is helpful to divide EDNOS in the categories atypical anorexia nervosa, atypical bulimia nervosa and binge eating disorder. The category mixed eating disorder is in my opinion rather unspecific and perhaps not very helpful."
22. "The "sleight of hand" to move a large proportion of EDNOS to "mixed eating disorder" still does not get around the problem of lack of positive criteria. Until research highlights the outer edges of caseness, it would still be largely down to clinicians'/researchers' discretion. Therefore, is this a useful exercise?"

Other comments:

23. "The NOS group is challenging the categorical way of subtyping disorders; it would be more useful to speak about core features ('typical' syndrome) versus variants, either qualitative or quantitative ('atypical' presentations); cf ICD approach. It is also known that research criteria are usually more strict than the ones clinicians use in their everyday practice! Whatever diagnostic system remains an instrument the value of which depends on its purpose and user..."
24. "Besides the DSM IV categorical diagnostic schemata, the problem of EDNOS arises from the DSM overvalue of behaviours with detriment of cognitions. I support the use of a dimensional approach investigating the core beliefs of ED patients."

4.4.5 Summary of findings

- A questionnaire survey was conducted assessing the opinion of eating disorder experts (attending a specialist meeting) on the problems associated with the DSM-IV diagnosis eating disorder NOS and the clinical utility of Fairburn and Bohn's 'Interim proposal' (2005) with respect to its validity, acceptability, ease of use and potential as a solution.
- One hundred and ninety questionnaires were sent out of which 68 were returned (35.8% response rate).
- The majority of the 68 respondents (between 73.6% and 92.7%) agreed with the problems associated with eating disorder NOS identified by Fairburn and Bohn (2005).

- A small proportion of respondents (14.7%) thought that relaxing the diagnostic criteria for anorexia nervosa and bulimia nervosa would be sufficient on its own to address the problems associated with eating disorder NOS.
- Three-quarters of the respondents (72.1%) thought that the addition of Step 2 of the 'Interim proposal' was needed (i.e., reclassifying the remaining cases as "mixed eating disorder" or BED).
- More than three-quarters of the respondents thought that the 'Interim proposal'

  had merit to it (77.9%),

  would draw attention to eating disorder NOS patients with regard to defining their clinical characteristics (76.5%),

  would encourage research on their treatment (80.9%),

  would be an improvement over the existing DSM-IV diagnostic scheme (82.3%),

  would not make the diagnostic system for eating disorders too complex (82.4%),

  would be usable in practice (77.9%).
- Over half the respondents thought that the 'Interim proposal'

  would address some or all of the problems associated with eating disorder NOS (60.3%),

  would contribute to defining the clinical needs of patients with eating disorder NOS (58.9%), and

  matched clinical reality as they see it in their clinical practice (69.1%).
- Less than half of respondents (45.6%) thought that it would be helpful for some of their eating disorder NOS patients to be given the diagnosis "mixed eating disorder".

- Additional comments from respondents showed a wide range of opinions with regard to the 'Interim proposal': about 60% of comments were generally positive, a third expressed doubts about its clinical usefulness, and the rest were more general. About half the comments were in favour of the 'Transdiagnostic solution' as the ultimate step forward in solving the problems of eating disorder NOS.

4.5 Discussion

Three solutions to the problem of eating disorder NOS

In an article published in *Behaviour Research and Therapy* Fairburn and Bohn (the candidate) (2005) highlighted certain problems associated with the DSM-IV diagnosis eating disorder NOS, namely the problems of nosology and neglect. They then proposed three solutions to these problems. The first solution was based on the premise that the high prevalence of eating disorder NOS cases is due to the DSM-IV diagnostic criteria for anorexia nervosa and bulimia nervosa being inappropriately strict. In involved systematically adjusting the diagnostic criteria for anorexia nervosa and bulimia nervosa in order to embrace their subthreshold cases currently within eating disorder NOS. The second solution was an elaboration and extension of the first. It involved two steps: the first was to adjust the diagnostic criteria of anorexia nervosa and bulimia nervosa as described in solution one, and the second was to reallocate the remaining cases (of which there were thought to be many) to a new diagnostic category, perhaps termed "mixed eating disorder" or to binge eating disorder. The third solution was to create a single unitary diagnostic category "eating disorder" embracing anorexia nervosa, bulimia nervosa and eating disorder NOS without any subdivisions. This transdiagnostic solution was based on the premise that the three eating disorders are far more similar than dissimilar.

Fairburn and Bohn argued that in the short term the second solution might have most

to recommend it since the first ignored the fact that many of the cases within eating disorder NOS are of the mixed variety (see Section 4.3 in this chapter). They acknowledged that the introduction of a new eating disorder diagnosis (possibly termed mixed eating disorder) was inconsistent with the conservative spirit of DSM-IV but, as Nielsen and Palmer have pointed out, "There is room for a measure of conservatism but we cannot be satisfied until the EDNOS issue is more adequately addressed" (Nielsen & Palmer, 2003, p. 162).

Fairburn and Bohn acknowledged that the re-classification of the cases within eating disorder NOS (inherent to solution two) was something of a sleight of hand, but they argued that it was a sleight of hand with a purpose since it placed these cases in specific and appropriate diagnostic categories which might enhance the credibility and usefulness of the scheme for classifying eating disorders and, hopefully, facilitate research on these problems including research on their treatment. They suggested that the proposal fulfilled some of First and colleagues' (2004) criteria for "clinical utility"

As regards Fairburn and Bohn's "transdiagnostic" solution, they suggested that in the longer term it might be the most valuable one. They suggested that the existing scheme for classifying eating disorders was a historical accident and that it was a poor reflection of current clinical reality. They thought that the transdiagnostic solution would encourage and permit the classification of eating disorders to be examined afresh. In particular it would encourage the collection of transdiagnostic data, particularly cross-diagnostic information on course and response to treatment, data of this type being needed for new clinically informative subdivisions to be identified.

The impact of adjusting the DSM-IV criteria for Anorexia nervosa and Bulimia nervosa on the relative prevalence of Eating Disorder NOS

The aim of this study was to determine whether the high proportion of eating disorder NOS cases within adult outpatient settings might be due in part to the presence

within them of cases closely resembling anorexia nervosa or bulimia nervosa, cases that might be better re-diagnosed as such. In other words this was a test of Fairburn and Bohn's (2005) first solution to the problem of the surfeit of eating disorder NOS cases; namely, removing subthreshold cases of anorexia nervosa and bulimia nervosa from the diagnosis eating disorder NOS. The availability of EDE ratings on a representative clinical sample of eating disorder cases made it possible to examine the effect of this solution on the relative prevalence of eating disorder NOS by systematically relaxing the diagnostic criteria for anorexia nervosa and bulimia nervosa. With regard to anorexia nervosa, the amenorrhoea criterion was dropped, the weight threshold was raised from 17.5 to a BMI of 18.0 and then to 18.5, and the core psychopathology was re-defined to include cases who show an over-evaluation of controlling eating per se. With regard to bulimia nervosa, the frequency threshold for objective bulimic episodes and compensatory behaviour was lowered from twice per week to on average once per week (over the past three months). Initially these modifications were made in isolation, and their impact on the relative prevalence rate of eating disorder NOS calculated. Then all modifications were done in combination, and their overall impact on the relative prevalence of eating disorder NOS was evaluated.

It emerged that none of the adjustments, either in isolation or in combination, had much impact on the relative prevalence of eating disorder NOS which remained at fifty percent or more. Nor was the presence of binge eating disorder responsible for the large number of eating disorder NOS cases for just seven patients (6.9% of the 102 eating disorder NOS cases) met its diagnostic criteria. [This figure rose to nine if the minimum average frequency of binge eating was reduced to once per week.] The low prevalence of binge eating disorder was not unexpected as three of the five patient samples referred to above (Section 4.1) reported prevalence figures for binge eating disorder and in all three instances

the prevalence was less than ten percent (Martin et al., 2000; Nollett & Button, 2005; Ricca et al., 2001).

Methodological considerations

A major strength of this study is that it is the first one to examine the impact of systematically modifying the diagnostic criteria of anorexia nervosa and bulimia nervosa on the prevalence rates of the three eating disorders, using a representative sample of adult outpatients. Two further strengths of the study have already been described in Chapter Three (Section 3.4), since the sample used in this study was identical to that used in Chapter Three. The first was the two-site catchment area sampling frame, which made it likely that the patients were representative of many other outpatient samples of adults with eating disorders. The second was that the cases were well characterised. The use of operational EDE-based diagnostic criteria enabled diagnostic thresholds to be adjusted in a systematic and replicable way.

Similarly, some of the limitations of the study have already been discussed (in Chapter Three, Section 3.4). Possibly the main one is that the sample did not include patients at the extremes of the weight spectrum (i.e., those with a BMI <16.0 or ≥ 40.0). Data from the Leicester clinic for the three-year period 2003 to 2005 indicate that these two groups constituted 8% (39/490) and 5% (23/490) respectively of the referrals (Palmer, 2006, personal communication). Had they been formally evaluated in the context of the present study patients with a BMI below 16 would have been likely to have fulfilled diagnostic criteria for anorexia nervosa whereas those with a BMI of 40 or more would have been likely to have been cases of eating disorder NOS. If these two groups had been included in the present sample the diagnostic distribution would have been 11% anorexia nervosa, 31 % bulimia nervosa and 58% eating disorder NOS, figures that differ little from those of the

actual study sample (4.7%, 35.3% and 60.0% respectively). The exclusion of these two small subgroups of patients will have had little effect on the study findings as these concern changes in prevalence rates following adjustments to the diagnostic criteria of bulimia nervosa and anorexia nervosa. Another potential limitation of the study concerns the adjustments that were made to the diagnostic criteria for bulimia nervosa: it was not examined how the relative prevalence rate of eating disorder NOS would change if the definition of a binge was expanded to include episodes of uncontrolled eating that do not involve eating a large amount of food (i.e., the inclusion of 'subjective bulimic episodes'). The reason for this was that it was decided to change the diagnostic criteria of anorexia nervosa and bulimia nervosa only in ways that would not alter the core clinical concepts of the two disorders. In the case of bulimia nervosa this meant preserving the requirement that patients with bulimia nervosa experience binges that are large in size. This requirement has however been questioned recently (Le Grange et al., 2006; Pratt et al., 1998; Rossiter & Agras, 1990), and it would certainly be worth examining the extent to which the inclusion of subjective bulimic episodes into the diagnostic criteria for bulimia nervosa would reduce the relative prevalence rate of eating disorder NOS.

Conclusions

Overall the findings from this study suggest that the high relative prevalence of eating disorder NOS is not attributable to the existence within it of cases closely resembling anorexia nervosa or bulimia nervosa or to the presence of cases of binge eating disorder. On the basis of this study's findings it can be concluded that adjusting the diagnostic criteria for anorexia nervosa and bulimia nervosa along lines advocated in the literature is not sufficient to solve the problems of nosology and neglect associated with the diagnosis eating disorder

NOS. Doing so in a representative sample of clinical eating disorder cases still left over half of the cases in the residual 'not otherwise specified' category.

Therefore it seems that only a small proportion of eating disorder NOS cases could be reasonably re-diagnosed as cases of anorexia nervosa or bulimia nervosa. The majority seem to be quite different whilst still having the psychopathology that characterises eating disorders. General clinical experience and knowledge of the cases that comprise the present sample suggest that most might best be described as "mixed" in character as the clinical features of anorexia nervosa and bulimia nervosa are present but combined in different ways to those seen in the two classic syndromes.

The practical implications of these findings are that more needs to be done than just relaxing the diagnostic thresholds for anorexia nervosa and bulimia nervosa if the eating disorder NOS problems of nosology and neglect are to be solved. Solving these problems requires changing the diagnostic system for eating disorders in a way that results in a) more attention being given to the considerable amount of patients currently being diagnosed with eating disorder NOS, and b) the 'shrinking' of the diagnostic category eating disorder NOS, so that it is a truly residual one. One way of solving the eating disorder NOS problems was suggested by Fairburn and Bohn (2005), in their so-called 'Interim proposal': after the expansion of the diagnostic criteria for anorexia nervosa and bulimia nervosa (as suggested above) the remaining cases of eating disorder NOS would be reallocated to binge eating disorder (although this study's findings suggest that there will only be few) or to a new diagnostic category, perhaps termed "mixed eating disorder". The clinical utility of this option was examined in Section 4.4 of this chapter.

Evaluating the clinical utility of Fairburn and Bohn's (2005) "Interim Proposal"

The aim of this study was to evaluate, by means of a survey, the opinion of eating disorder experts on: a) the DSM-IV diagnosis eating disorder NOS and its associated problems of nosology and neglect; b) the 'Interim proposal' suggested by Fairburn and Bohn (2005) designed to solve these problems; c) the potential of this proposal for addressing the problems associated with eating disorder NOS; and d) the diagnostic validity of the proposed change as well as their acceptability and ease of use. This was done in order to assess one aspect of clinical utility, as described by First and colleagues (2004), of the proposed change to the diagnostic system.

The majority of the respondents agreed with the problems associated with eating disorder NOS identified by Fairburn and Bohn (2005). Most thought the proposed change had merit and felt that Step 1 of the 'Interim proposal' (i.e., adjusting the diagnostic criteria of anorexia nervosa and bulimia nervosa) would not be sufficient on its own to solve the problem of eating disorder NOS, but that Step 2 was needed. As indicated by the findings of the second study in this chapter (Section 4.3) they were correct in this view since Step 1 on its own has little effect on the relative surfeit of eating disorder NOS cases.

Between 60% and 80% of the respondents were of the opinion that the proposed change would draw attention to eating disorder NOS patients with regard to defining their clinical characteristics and needs, and encouraging research on their treatment. Over two-thirds thought that the proposed change matched clinical reality as they see it in their clinical practice; three-quarters thought that it would be usable in practice; over eighty percent thought that it would not make the diagnostic system for eating disorders too complex and would be an improvement over the existing DSM-IV diagnostic scheme for eating disorders.

On the other hand, less than half thought that it would be helpful for some of their eating disorder NOS patients to be given the diagnosis "mixed eating disorder".

The additional comments that respondents provided were neither uniformly positive or negative with regard to the proposed changes. About 60% of the comments were generally positive, a third expressed doubts about its clinical usefulness, and the rest were more general. It was interesting to see that nearly all of the respondents that commented positively about the 'Interim proposal' pointed out that it was a good short-term solution but that further steps would need to be taken to efficiently deal with the problems of eating disorder NOS. Fifty per cent of the comments were in favour of the 'Transdiagnostic solution' as the ultimate step forward in solving the problems of eating disorder NOS.

Methodological considerations

This study has some important limitations. The first concerns the sample. Although the sampling frame was a good one in that is was composed of informed clinicians, it is not clear that they were representative of those who use DSM (Diagnostic and Statistical Manual for Mental Disorders). It was also small in size. Furthermore, it may have been biased in that only 35.8% of the eligible participants returned the questionnaire. It is not possible to predict the nature of the bias, however, since the comments received indicated that they were neither uniformly positive or negative regarding Fairburn and Bohn's (2005) proposed solutions.

The questionnaire also had shortcomings. It was decided to use a self-report questionnaire. An interview would have been preferable since it would have been likely to have provided more detailed information. However, an interview was not practicable under the circumstances since over fifty busy conference attendees would have had to been interviewed in a very short time. Also, the Likert scoring scheme was flawed. As one

participant commented, there was only one "disagree" rating option on the four-point scale (and it was phrased strongly), whereas there were four positive ratings. It is possible that some respondents did not disagree 'strongly' with certain items but at the same time did not agree with them. It would have been difficult for such respondents to choose a rating that accurately represented their views.

Conclusions

Overall the findings of this study suggest that eating disorder NOS is viewed as a problematic eating disorder category by expert eating disorder clinicians. There also seems to be consensus that it would not be enough to just relax the diagnostic criteria for anorexia nervosa and bulimia for the problems of eating disorder NOS to be solved, namely the high prevalence of the diagnosis in most adult outpatient settings and the neglect of patients with the disorder with regard to treatment and research. Most of the expert eating disorder clinicians appeared to be keen for the diagnostic scheme for eating disorders to be changed. There were a variety of opinions regarding how exactly this should be done. Perhaps surprisingly, a substantial subgroup was in favour of Fairburn and Bohn's most radical solution, the 'transdiagnostic' one.

As pointed out by a number of respondents, whatever the step forward will be with regard to improving the diagnostic scheme for eating disorders, it will be of ultimate importance to define the outer boundary of what is an eating disorder (see Figure 1.1, Section 4.2.1). If a new category of eating disorder is created, there will have to be positive criteria for it. If one diagnostic category called 'eating disorder' is created (in order to then study meaningful subgroups within it) it will have to be delineated from lesser, non-clinical eating problems, i.e., there will have to be a definition of what is an eating disorder. This topic is addressed in Chapter Five.

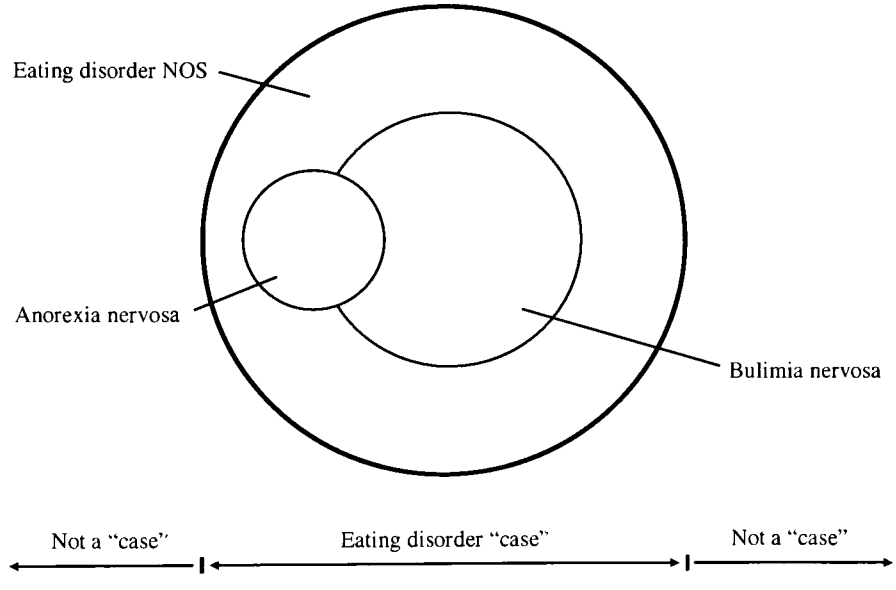
CHAPTER FIVE

Defining the Boundaries of an "Eating Disorder"

5.1 Introduction

As discussed in Chapter One (section 1.3.2), there is currently no officially agreed definition of what constitutes an 'eating disorder'. Diagnostic criteria are specified in DSM-IV for the two eating disorders anorexia nervosa and bulimia nervosa but these have been criticised for being arbitrary and inappropriate. No diagnostic criteria are specified for the most common eating disorder, eating disorder NOS, and therefore it is not clear what constitutes an eating disorder "case" (as illustrated in Figure 1.1 (taken from Chapter One)).

Figure 1.1 *A schematic representation of the relationship between anorexia nervosa, bulimia nervosa and eating disorder NOS*



Defining what is an eating disorder (i.e., the "outer edge" of Figure 1.1 is possible. Given the definition of a psychiatric disorder (see Chapter One, section 1.3.3), it requires the identification of the type and level of eating disorder features that are associated with

clinically significant impairment of functioning. Doing so, however, is at present limited by the absence of a measure of impairment secondary to eating disorder psychopathology. In this Chapter the development of such a measure is described (Section 5.2). Studies of the instrument's validity and reliability are presented in Section 5.3. Finally, this instrument is used to derive a provisional impairment-based operational definition of what constitutes an eating disorder (Section 5.4). The findings of all these studies are summarised at the end of each section and discussed in section 5.5.

5.2 Development of the Clinical Impairment Assessment Interview (CIA)

In this section of the chapter the major steps in the development of the Clinical Impairment Assessment interview (CIA) are described. This instrument was developed by the candidate and three experts in the eating disorder field: Professor Christopher Fairburn, Dr Zafra Cooper and Professor Robert Palmer. The preliminary version of the CIA, and the various versions of the CIA proper, are summarised, and a detailed account of the current version of the CIA (CIA 3.0) is given. The current version is reproduced in Appendix 5.1.

The first step in the development of the CIA was to decide what basic features the instrument required. Given its purpose, it was decided that it should have the following characteristics:

- Investigator-based, semi-structured interview - to ensure that the secondary psychosocial and physical impairment is explored in an individualised and detailed way;
- Should focus exclusively on domains of life that are affected by eating disorder features – to ensure the relevance of its content;

- Should focus on the participant's state over the past 28 days – to ensure that it assesses current impairment;
- Should have a rating scheme that distinguishes clinically significant levels of secondary impairment from milder non-clinical levels;
- Should be administered immediately after the assessment of current eating disorder psychopathology (using an interview such as the EDE) – to ensure that any eating disorder features present are brought to the forefront of the participant's mind before he or she answers questions about how these features have been interfering with his or her life;
- Should help participants consider all aspects of their eating disorder psychopathology when evaluating how their life has been affected – to ensure that the full extent of any secondary impairment is rated;
- Should ensure that all impairment rated is secondary to eating disorder features per se and not due to other causes (e.g., comorbid psychiatric conditions, life difficulties).

5.2.1 The content of the CIA

The second step in the development of the CIA involved thinking about the content of the instrument, i.e., identifying the domains of life most prone to be affected by eating disorder features. This was done in collaboration with the three eating disorder experts (Professor Christopher Fairburn, Dr Zafra Cooper and Professor Robert Palmer), each of whom had been working in this field for at least two decades and had extensive experience assessing and treating patients with eating disorders. All three also had a particular interest in the impact an eating disorder has on sufferers' lives. The candidate herself contributed to

this process by using her own clinical experience[7] and by asking eating disorder patients directly what impact their eating disorder was having on their own lives. It was concluded that the following five domains of life were most prone to be affected by eating disorder features.

1. **Mood and Self-perception** – i.e., one's feelings and the way one views oneself as a person;
2. **Cognitive functioning** – i.e., one's thinking and ability to concentrate;
3. **Interpersonal functioning** – i.e., one's relationships and the way one gets on with others;
4. **Work performance** – i.e., the extent to which one is able to perform work or household tasks;
5. **Physical health.**

The third step in developing the CIA involved specifying real-life examples of impairment in each of the five identified domains. This was also done by the candidate in collaboration with the three expert clinicians. As in the process of identifying the individual domains, clinical experience and patients' accounts of their impairment were used to identify and specify examples of secondary impairment. These are shown in Table 5.1.

Table 5.1 *Real-life examples of impairment due to eating disorder features in five domains of life*

CIA domains	Examples
Mood and Self-Perception	• Self-criticism over the way one is eating or one's shape/weight • Distress when not able to prepare own food • Irritability at meal times • Feeling guilty about having eaten • Distress over seeing weight on scales • Distress over having broken dietary rules
Cognitive functioning	• Preoccupation with thoughts about food/eating • Difficulty concentrating due to thoughts about shape and weight • Being unable to concentrate before and after binge eating • Not being able to focus on conversations due to thoughts about eating • Being significantly forgetful and absent-minded • Difficulty making everyday decisions due to preoccupation with shape

[7] The candidate was a research therapist at the Centre for Research in Eating Disorders in Oxford from October 2001 to September 2006. During this time she treated more than 50 patients with clinical eating disorders and thereby gained an intimate knowledge of these patients' problems.

Interpersonal functioning	• Decreased socialising due to drive to exercise/fear of overeating/shame about shape • Arguments with family/partner over eating habits • Inability to eat with others • Impairment of sex life
Work Performance	• Being distracted from work by feeling need to check body • Taking longer to complete tasks, but no adverse consequences • Repeatedly having to leave work in order to buy food/binge/vomit • Being reprimanded for poor work performance due to disturbed eating • Having to take a year off from college due to eating disorder
Physical health	• Being significantly underweight • Amenorrhoea • Electrolyte disturbance due to frequent vomiting

5.2.2 The rating scheme of the CIA

The rating scheme for the five domains

The next step was to create a rating scheme for each life domain, reflecting the severity of secondary impairment present. On the basis of the list of examples of impairment for each life domain and clinical consensus amongst the three expert clinicians and the candidate, it was decided to use a three-point rating scheme (0-2) for each life domain, distinguishing between no, occasional/mild, and definite (clinically significant) secondary negative effects. It was agreed that a rating of 2 (clinically significant impairment) should only be given, if the secondary impairment was:

- Moderate in severity, but occurring frequently, and/or
- Marked in severity, and occurring regularly (at least once a week), and/or
- Of a quality that was clearly significantly interfering with the person's life.

The individual ratings and their definitions are shown in Table 5.2.

Table 5.2 CIA ratings for each life domain and their operational definitions

Rating of the severity of impairment in each domain	Extent of secondary effects on functioning in respective domain	Operational definitions of extent of secondary effects
0	No secondary effects	
1	Some secondary effects (occasional and/or mild)	Persistent low-grade or occasional marked effects (i.e., less than four times over the past month)
2	Secondary effects of clinical significance	Persistent moderate or regular marked effects (i.e., four or more times over the past month)

After the rating scheme for the domains had been created, examples of real-life impairment were assigned to the three severity ratings for each life domain in order to give interviewers guidance as to how to rate any secondary impairment elicited. Examples for a rating of 2 in each life domain are shown in Table 5.3 below. The thresholds between ratings of 1 and 2 were derived on the basis of clinical consensus amongst the three clinicians and the candidate.

The global impairment rating

It was also agreed by the candidate and the three clinicians, that, if a participant had a rating of 2 in any of the five domains of life, their overall secondary impairment (i.e., global impairment) would be regarded as "clinically significant". In addition, it was decided that if a person showed some degree of secondary impairment (i.e., a rating of 1) across three or more of the five domains (i.e., there was pervasive impairment), their impairment would also be classed as clinically significant. In other words, the level of global impairment present was to be viewed as clinically significant if there was either severe impairment in one domain of life or less severe but pervasive impairment across three or more life domains.

Table 5.3 Examples of clinically significant impairment (a rating of '2') in each domain of life

CIA domains	Examples of clinically significant impairment (i.e., a rating of '2')
Mood and Self-Perception	• Pronounced and persistent self-criticism over inability to control eating or one's shape and/or weight • Continuously worrying about being able to exercise enough • Marked irritability at most evening meals • Distress over binge eating (on most days)
Cognitive functioning	• Frequently not being able to focus on conversations/interesting TV programs/reading due to thoughts about eating, shape or weight • Being significantly forgetful and absent-minded • Difficulty making everyday decisions due to preoccupation with shape/weight
Interpersonal functioning	• Persistent interpersonal tension over the way the person eats • Significantly decreased socialising due to drive to exercise/fear of overeating/shame about shape • Frequent arguments with family/partner over eating habits • Inability to eat with others • Significant impairment of sex life (i.e., not wanting to be touched) • Not wanting to enter a relationship because it will be difficult to continue vomiting
Work Performance	• Repeatedly having to leave work in order to buy food/binge/vomit • Being reprimanded for poor work performance due to disturbed eating • Having to take a year off from college due to eating disorder
Physical health	• Being significantly underweight (body mass index below 18.0) due to undereating • Amenorrhoea • Electrolyte disturbance due to frequent vomiting (twice or more daily)

5.2.3 The style of the CIA

The fifth step in the development of the CIA was to compose questions that would elicit the impairment present in each life domain. Given that it had been decided that the CIA should be an investigator-based interview, the questioning style of the Eating Disorder Examination (EDE) was adopted (Fairburn & Cooper, 1993). Thus, it was decided that each life domain should be entered with a *mandatory probe question* asking the participant whether this particular domain of life (e.g., their relationships and the way they get on with others) had been affected by (specified) eating disorder features over the previous four

weeks. Each domain was also provided with a series of *mandatory additional questions*, their purpose being to elicit further relevant information and examples of any secondary impairment present. It was also made explicit in the instructions for interviewers that both the frequency and intensity of any impairment present should be assessed. Once all mandatory questions had been asked, the interviewer was encouraged to ask more questions as needed to obtain further detail about the impairment in question. Figure 5.1 shows the questions for the first of the five domains, viz., "Mood and Self-Perception", taken from the latest version of the CIA (CIA 3.0; see Appendix 5.1).

Figure 5.1 *The domain "Mood and Self-Perception" from the CIA, version 3.0*

DOMAIN 1 - MOOD AND SELF-PERCEPTION

Mandatory probe question:

*** Over the past four weeks have** ...(refer to participant's eating disorder features)... **affected you emotionally?**.....

Mandatory additional questions: (Obtain clear examples and elicit details of frequency and severity.)

*** Have they affected your mood?**
*** Have they upset you?**
*** Have they made you feel depressed or irritable or worry a lot?**
*** Have they affected how you view yourself?**
*** Have they made you feel critical of yourself?**
*** Have they made you feel guilty - made you think that you have done wrong?**

[Rate secondary effects. Ensure that any impairment is a result of eating disorder features.]

0 - No secondary effects on mood or self-perception

1 - Some secondary effects (occasional and/or mild). These may comprise persistent low grade effects or occasional marked disturbance *(on less than four occasions over the past four weeks).*

- continuous mild self-criticism over the way one is eating
- continuous mild self-criticism over one's weight and/or shape
- mild distress over not being able to adhere to dietary rules
- distress when occasionally not able to prepare own food
- distress (less than once per week) due to fear of losing control over eating
- feeling less self-confident on days when has binged
- distress confined to occasional binge eating (less than once a week)
- "disliking" the fact of self-induced vomiting
- distress over seeing weight on scales, but feelings do not last long
- distress over clothes feeling too tight (intermittent moderate?)
- feeling unhappy whenever getting dressed

2 - Definite secondary effects. These may comprise persistent moderate effects or regular marked effects *(on at least four occasions over the past four weeks).*

- distress every evening over having to break dietary rules (due to having to eat with others)
- being very self-critical due to feeling "compelled" to restrict eating

- pronounced and persistent self-criticism over inability to control eating
- continuously worrying about being able to exercise enough (in context of "driven exercising")
- pronounced and persistent self-criticism over one's shape and/or weight
- marked irritability at most evening meals
- distress over binge eating (on most days)
- disgust on seeing one's body
- being highly self-critical whenever thinking about one's shape
- guilt over eating due to being a bad role model for children (children might pick up one's bad habits)

7 - Not applicable

[Rate using the following time frames (as needed) in the order listed below).]

last 4 weeks [ ]; last 3 months [ ]

5.2.4 The use of the CIA

Once the content and rating scheme of the CIA had been determined, it was agreed that it should be administered immediately after a thorough assessment of the current eating disorder psychopathology (using the EDE). An important matter to be decided was what level of eating disorder features would need to be present for an assessment of secondary impairment to be relevant. As a default it was decided that the CIA should be administered unless the participant had few or no behavioural features of an eating disorder (i.e., successful dietary restraint or restriction, objective or subjective bulimic episodes, self-induced vomiting, laxative or diuretic misuse). A low threshold for entering the instrument was deliberately set so that impairment due to all levels of eating disorder psychopathology could be assessed. If the interviewer was in doubt about whether to administer the instrument (i.e., they were not sure whether the eating disorder features were severe enough to merit entering the CIA), the instruction was to proceed with the CIA and seek advice from Professor Fairburn later as to whether this had been necessary or not. In version 2.1 of the CIA (see section 5.2.5 of this chapter) detailed (EDE-) thresholds were introduced regarding what constituted "non-trivial" levels of eating disorder psychopathology (i.e., the level of eating disorder psychopathology that merited entering the CIA).

The interviewers received extensive training by the candidate in how to administer the interview with each new version that was produced. They were asked to rate each life domain as they proceeded with the interview and to end the interview by giving a rating of 0 or 1 for the participant's global impairment (following the rules specified above), 1 being "clinically significant global impairment". They were also asked to review their ratings (for each life domain and the global rating) with either the candidate or Professor Fairburn as soon as possible after the interview. This was done without the interviewer revealing the identity of the participant in question. In addition, the interviewers had regular meetings amongst each other during which they discussed participants' CIAs in order to 'cross-calibrate' their ratings.

5.2.5 The CIA – from the preliminary to the current version

A preliminary version of the CIA, four early versions and the current version of the CIA exist. From the outset, each version of the CIA was administered as part of the assessment of the patients with a clinical eating disorder described in Chapter Three. These patients were consecutive referrals to two eating disorder clinics, one in Oxford and one in Leicester (see Chapter Three, Section 3.2 for details). Routine research assessments took place before and after 20 or 40 weeks of cognitive behaviour therapy; and at 20, 40 and 60 weeks and two and three years post-treatment. Administering each version of the CIA to patients with all levels of eating disorder psychopathology was crucial for refining the instrument and for developing the current version. In the section below, the preliminary version, four early versions and the current version of the CIA are described. It is also pointed out what changes were made from one version to the next. It is important to note that the data for all the studies presented later in this chapter were acquired using either CIA version 2.0, 2.1 or 3.0 (the current version).

Preliminary version of the CIA

CIA version P.1

The preliminary version of the CIA was administered over a period of five months (February to July 2003) as part of the routine assessment at the beginning of treatment.

The interviewers were asked to enter the instrument immediately after having administered the EDE saying the following to the participant:

> *"I now would like to ask you how the things we have just been talking about, i.e., the way you have been eating and the way you have been feeling about your eating, shape or weight, have been influencing your life over the past month."*

The interviewers then entered the sections of the CIA that addressed the five domains of life, the form of the interview resembling the current version (see Appendix 5.1).

The CIA proper

CIA version 1.0

This version of the CIA was used for six months between August 2003 and January 2004. It was administered at the following assessment time points: beginning and end of treatment, and at 20-, 40-, and 60-week post-treatment. It differed from the preliminary version of the CIA in the following ways:

1. The CIA was entered with the interviewer summarising the participant's individual eating disorder features as assessed with the EDE (e.g., objective bulimic episodes, self-induced vomiting, following strict dietary rules).

   This change was made in order not to miss out secondary impairment that might be due to a specific eating disorder feature. This might have not been detected by the previous version of the CIA in which participants were asked a much more general

question about impairment due to 'the way that they have been eating or feeling about their eating, shape or weight'.

2. A sixth life domain was added, termed "*Life not otherwise specified*" ("Life NOS"). This was done because it had become apparent that a number of participants reported impairment that was un-rateable in the preliminary version of the CIA since it did not fit into any of the existing domains. Examples of this type of impairment were financial difficulties/debts due to buying food for binges, and not being able to engage in a hobby one used to enjoy (e.g., swimming, horse riding) because of shape concerns or preoccupation with shape and weight.
3. Certain additional mandatory questions were added to the instrument in order to elicit secondary impairment in patients of the following types:
    - Patients who experienced having an eating disorder as largely positive (i.e., they felt good about controlling their eating),
    - Patients who were ambivalent about overcoming their eating disorder and therefore denied any secondary impairment,
    - Patients who had had their eating disorder for a long time and therefore had 'accommodated' it in their every-day life (i.e., they were not aware of any impairment present, as it felt normal to live their life like this).

    "Over the past four weeks would your life have been significantly better in any way if:
    .... you had been able to eat like others (..... without an eating problem);
    ..... in other words you had been able to eat without (..... refer to the patient's dieting/vomiting/laxative misuse/binge eating)?"

If the answer to the above question was "No", the following question was asked:

> *"Over the past four weeks would your life have been significantly better in any way if you had been less unhappy about your eating, shape or weight?"*

If the answer to either of these two questions was "Yes", the interviewer went on to assess in detail how the participant's life would have been better. Any secondary impairment detected in this way was subsequently rated in the relevant life domain.

CIA version 1.1

This version of the CIA was used for four months (January to May 2004). It was administered at the following assessment time points: beginning and end of treatment, and at 20-, 40-, and 60-week post-treatment follow-up.

The following new features were introduced:

1. An additional question was added to the 'Interpersonal impairment' domain. If the participant had not had the opportunity to socialise or eat with others in the previous 28 days, interviewers were asked to explore whether the participant *"could have done"* either of the above.
2. Further examples of impairment that had been elicited with previous versions of the CIA were added for ratings of 1 and 2 in each domain.

CIA version 2.0

This version of the CIA was administered at all assessment time points over a time period of six months (May to November 2004).

In this version a significant change was made to the way the CIA was entered. After having completed the EDE interview, interviewers were asked to establish which eating disorder features were present considering all positive EDE items (i.e., behavioural features as well as shape-and weight-related items). This involved going over the EDE scoring sheet

and reviewing all the eating disorder features that were rated as being present and then checking back with the patient. The goal was to create a comprehensive list of the main eating disorder features present in rough order of severity; for example:

1. *Extreme fear of weight gain*
2. *Dissatisfaction with shape*
3. *Regular binge eating*
4. *Difficulty eating in front of others*
5. *Having to follow multiple dietary rules*
6. *Possibly eating too little*

The interviewers were asked to take the lead in creating this list. They then entered the CIA domains, having the list in front of them and the participant, and they then referred to the list when assessing any impairment.

This change was made to ensure that impairment due to all eating disorder features present was assessed, and that the assessment was consistent across all six domains of life.

CIA version 2.1

This version of the CIA was administered at all assessment time points over seven months (November 2004 to July 2005). It differed from version 2.0 in the following way. On the front page of the CIA guidelines were provided about what constituted 'non-trivial' levels of behavioural eating disorder features (over a time period of 28 days). This was done to help interviewers decide whether to administer the CIA or not. Interviewers were asked to proceed with the interview if the person in question had shown one or more of these features at a non-trivial level over the previous 28 days. The non-trivial levels were derived on the basis of clinical consensus amongst the three eating disorder clinicians and the candidate. They are shown in Table 5.4.

Table 5.4 *Guidance for interviewers on the front sheet of the CIA regarding "non-trivial" levels of behavioural eating disorder features (over a time period of 28 days)*

EDE items	**"Non-trivial" ratings / frequency of episodes (as assessed with the EDE)**
Restraint over eating – only rate adherence	≥ 2
Avoidance of eating	≥ 2
Food avoidance – only rate adherence	≥ 2
Dietary rules – only rate adherence	≥ 2
Objective bulimic episodes (OBEs)	≥ 2
Subjective bulimic episodes (SBEs)	≥ 4
Eating in secret	≥ 2
Self-induced vomiting	≥ 2
Laxative or diuretic misuse	≥ 2
BMI	< 18.5
Actual dietary restriction (i.e., significant undereating) → look out for infrequent or delayed eating when assessing the patient's temporal pattern of eating	see left

CIA version 3.0 – The current version

The current version of the CIA has been administered at all assessment points since July 2005. It can be seen in Appendix 5.1. The only difference between the current version and version 2.1 lies in its layout. Previously, all domain ratings and the global impairment rating were made on the interview schedule itself. With this version, a separate CIA coding sheet was used (see Appendix 5.2).

5.3 The Reliability and Validity of the CIA

Three studies were conducted to test the validity of the CIA and its performance. The aim of the first was to test its inter-rater reliability. The aim of the second was to establish its test-retest reliability. The aim of the third was to establish its construct and the

discriminant validity. The findings of each of these studies are summarised at the end of each study. The studies as a whole are discussed in the overall discussion in section 5.5.

The participants for all three studies were subgroups of the 170 patients described in Chapter Three. As outlined in Chapter Three, these patients were consecutive referrals to one of two outpatient eating disorder clinics in the UK (Oxford or Leicester) for the treatment of an eating disorder. The diagnostic distribution of the total sample at the time of referral was as follows: eight patients (4.7%) fulfilled diagnostic criteria for anorexia nervosa, 60 (35.3%) were diagnosed with bulimia nervosa, and 102 (60.0%) received the diagnosis eating disorder NOS.

As part of the research assessments at the beginning and end of their treatment, and at 20-, 40-, and 60-week post-treatment follow-up, and at two and three years after treatment, participants were interviewed with the Eating Disorder Examination (EDE; Fairburn & Cooper, 1993). They were subsequently interviewed with the CIA if they reported 'non-trivial' levels of disturbance of eating (see Table 5.4, Section 5.2.5 above) over the previous 28 days. The research interviews were conducted by one of six trained research assistants (Marianne O'Connor, Caroline Plumb and Shani Langdon in Oxford; Jackie Wales, Claire Nollett and Elizabeth Benson in Leicester). The patients were also seen by one of the three local expert clinicians (Dr Zafra Cooper or Professor Christopher Fairburn in Oxford; Professor Robert Palmer in Leicester) at each of these times, who made a clinical judgement as to whether the participant had an eating disorder of clinical severity. In Oxford, the two expert clinicians additionally formally assessed to what extent the participant's eating habits and concerns about their eating, shape or weight had influenced their psychosocial functioning over the previous 28 days (i.e., the same time frame as the one used in the CIA) (details of this assessment are below). Since this latter part of the

research assessment was only conducted in Oxford, none of the Leicester patients are included in the samples used for the test of construct validity or the test-retest reliability study.

The data of participants diagnosed by Professor Fairburn or Professor Palmer as suffering from a severe comorbid clinical depression were excluded since some of the impairment reported by such patients might have been secondary to the clinical depression rather than the eating disorder.

5.3.1 Study One: Inter-rater Reliability of the CIA

Design

The aim of this study was to evaluate the inter-rater reliability of the CIA, that is whether scores on the CIA are consistent when rated by two independent, trained assessors. The design involved the administration of the EDE and the CIA to a group of subjects with varying degrees of eating disorder psychopathology (i.e., exhibiting the complete spectrum of severity of eating disorder disturbance). These interviews were conducted by one of the six trained research assistants and were audio-taped. Subsequently the recordings of the interviews were listened to and rated by an independent assessor (the candidate). The ratings for each CIA domain and the global impairment ratings were compared. It was predicted that, if the CIA had good inter-rater reliability, there should be statistically significant agreement between the two independent raters on the CIA domain ratings and the global impairment ratings.

Participants

Thirty-one interviews were assessed for inter-rater reliability. The participants were 31 of the 170 patients described in Chapter Three. In order to have a group of participants with varying degrees of eating disorder psychopathology, it was decided to choose an equal

number of participants from the beginning of treatment (N=10), the end of treatment (N=10), and from one of the five follow-up assessments (N=11). Tape recordings were selected (by the candidate) from each of the three time frames blind to the participants' identity. A research assistant ensured that none of the tapes chosen were assessments of patients that the candidate herself had treated. Twenty-one of the 31 selected participants (67.7%) had an eating disorder of clinical severity at the time point of assessment. The remaining 10 (32.3%) no longer had a clinical eating disorder but most had residual eating disorder features.

Procedure

The 31 participants had been administered the EDE (Fairburn & Cooper, 1993) followed by the Clinical Impairment Assessment (CIA) by a trained research assistant as part of their research assessment. All the interviews were audio-taped. The tape recordings of the administration of the EDE and the CIA were subsequently listened to and rated by the candidate who was blind to the identity of the patient and the previous ratings.

Measures

The two measures in this study, the EDE and the CIA, have been described above (EDE in Section 3.2.2 of Chapter Three; CIA in Section 5.2.5 of this chapter). As explained above, only the latest three versions of the CIA (2.0, 2.1 or 3.0) were used.

Data analysis

The inter-rater reliability of the CIA was assessed in two ways. First, it was determined whether the three-level categorisation (0, 1 and 2) for impairment in each individual CIA life domain was stable across raters, and second it was determined whether the two-level global impairment rating (0 and 1) was also reliable. For both analyses the percentage of agreements between the two raters were calculated. In addition, Cohen's

kappa coefficient (Cohen, 1960) was calculated. Kappa is a chance-corrected measure of agreement, as it represents the proportion of agreement obtained after removing the proportion of agreement that could be expected to occur by chance. Kappa is always less than or equal to 1. A value of 1 implies perfect agreement whereas values of less than 1 imply less than perfect agreement. Kappa coefficients ranging from 0.40 to 0.59 are considered moderate, 0.60 to 0.79 substantial, and ≥0.80 outstanding (Landis & Koch, 1977).

Results

Table 5.5 shows the number and percentage of participants who were given an identical rating by the two raters in the CIA life domains and for the global impairment rating. Kappa coefficients and their statistical significance are reported where they could be calculated.

Table 5.5 *Inter-rater reliability of the CIA: Number and percentage of exact agreements between the two raters on the six CIA domains and the global impairment rating and kappa coefficients*

CIA domain	Exact agreement (N, (%))	Kappa coefficient
Mood and Self-Perception	26/31 (83.9%)	0.582***
Cognitive Effects	27/31 (87.1%)	0.801***
Interpersonal Impairment	23/31 (74.2%)	0.601***
Work Impairment	30/31 (96.8%)	0.942***
Life NOS	28/31 (90.3%)	0.689***
Physical Impairment	30/31 (96.8%)	
Global Impairment Rating	27/31 (87.1%)	0.527**

** significant on 0.01 level *** significant on 0.001 level

As can be seen, the percentage of exact agreement between the two raters was high across all life domains and with regard to the global impairment rating. All the kappa coefficients were statistically significant at least at a p=0.01 level, indicating a significant deviation from a kappa value of 0 (no agreement). According to the definition of kappa values by Landis and Koch (1977) agreement was "outstanding" for the domains Cognitive Effects, Work

Impairment, and Physical Impairment; "substantial" for the domains Interpersonal Impairment and Life NOS; and "moderate" for the domain Mood and Self-Perception and the global impairment rating.

Summary of findings

- The global impairment rating on the CIA (i.e., whether or not a participant was classed as being clinically significantly impaired by their eating disorder symptoms) was rated identically by two independent raters in 27 out of 31 participants (87.1%). The kappa value indicated that the agreement between the two raters regarding the global impairment rating was "moderate".
- The percentage of exact agreement in the ratings of the two independent raters was high for all six CIA life domains, ranging from 83.9% to 96.8%.
- There was a substantial-to-outstanding level of agreement between the two raters with regard to their ratings on five of the six life domains. Only the ratings on the Mood and Self-Perception domain showed a lower "moderate" level of agreement.
- The ratings of the second rater (the candidate) were based on tape recordings of the original interview, conducted by a first rater. This study therefore only tests one aspect of inter-rater reliability: the agreement in the way participants' responses are rated (see Discussion, section 5.5 of this chapter).

5.3.2 Study Two: Test-retest Reliability of the CIA

Design

The aim of this study was to evaluate the test-retest reliability of the CIA, that is the stability of ratings on the CIA over time, although the design also involved the use of different raters so that it was truly a mix of test-retest reliability and inter-rater reliability (see Discussion, section 5.5 of this chapter). The design involved administering the EDE

and the CIA twice within seven days to a group of participants with varying degrees of eating disorder psychopathology (i.e., exhibiting the complete spectrum of severity of eating disorder disturbance) and comparing both the CIA ratings for each life domain and the global impairment ratings. On each occasion the same 28-day time period was assessed. It was predicted that, if the ratings of the CIA were stable over time, there should be a statistically significant positive association between the CIA ratings at the two time points.

Participants

The sample comprised 11 of the 105 patients at the Oxford site described in Chapter Three. The data for this study were acquired between January 2004 and December 2005 and were from patients who were attending either their end-of-treatment or one of their post-treatment assessments[8]. Each participant was invited to participate in this study if they:

1. Had been interviewed with the CIA at the respective assessment time,
2. Had not received treatment from the candidate, and
3. Had not suffered from a comorbid clinical depression in the month before attending their assessment[9].

The candidate explained to the participant that the study was a research project that was entirely separate from the routine research assessment that they had just undergone and that taking part in it was voluntary. Twenty-six patients were eligible for the inclusion in this study. Nine of them no longer lived in Oxford at the time of assessment. Taking part would therefore have been impractical for them. Four of the 26 patients declined to take part. Two patients agreed to take part but did not attend the second assessment.

[8] These time points were chosen as it was judged to be too demanding for patients to go through an EDE and CIA assessment twice before starting treatment for their eating disorder.
[9] This was assessed by Professor Fairburn in a clinical assessment that took place at each assessment time point.

The sample for this study therefore comprised 11 patients. They took part at the following times: two at the end of treatment, two at 20-week follow-up, three at 40-week follow-up, two at 60-week follow-up, one at two-year follow-up, and one at three-year follow-up. Four of the 11 participants (36.4%) had an eating disorder of clinical severity (as judged by one of the two senior clinicians in Oxford) at the time of assessment.

Procedure

At the first assessment (T1), participants were administered the EDE followed by the CIA by a trained research assistant as part of their routine research assessment. The second interview (T2) took place three to seven days after T1 and also consisted of the administration of the EDE and the CIA. The interview at T2 was conducted by the candidate, who had been trained in administering the EDE. The candidate had no knowledge of the content of the interview at T1 and was thus blind to the previous ratings. At T2 participants were asked to consider the 28 days prior to T1 (i.e., not considering the intervening three to seven days) when answering the questions. This was done so that participants had the same period in mind as at the original assessment.

Measures

The two measures in this study, the EDE and the CIA, have been described above (EDE in Section 3.2.2 of Chapter Three; CIA in section 5.2.5 of this chapter).

Data analysis

The data analysis and statistical methods used for this study were identical to those described in Study One (5.3.1 Inter-rater reliability of the CIA). The only difference was that the sample size was not large enough to examine the test-retest reliability of the three-level categorisation of each CIA life domain (i.e., 0 versus 1 versus 2). It was therefore decided to collapse ratings 0 and 1 into one single rating (representing subthreshold levels of

impairment) and examine the test-retest reliability of each CIA life domain for the two resulting levels (0 and 1 vs. 2; i.e., subthreshold impairment versus clinically significant impairment). The temporal stability of the global impairment rating was examined as described in the previous study.

Results

Table 5.6 shows the percentage of exact agreements between the two raters on the CIA domains and the global impairment ratings at T1 and T2. Kappa coefficients and their statistical significance are also reported. Since not all possible ratings in the domains Work Impairment and Life Not Otherwise Specified were used by the assessors at T1 and T2, kappa coefficients could not be calculated for these two domains. The same was true for the global impairment rating.

Table 5.6 *Test-Retest reliability of the CIA: Number and percentage of exact agreements between the two raters on the six CIA domains and the global impairment rating and kappa coefficients*

CIA domain	Exact agreement (N, (%))	Kappa coefficient
Mood and Self-Perception	7/11 (63.6%)	0.35
Cognitive Effects	10/11 (90.0%)	0.62*
Interpersonal Impairment	9/11 (81.8%)	0.39
Work Impairment	10/11 (90.0%)	
Life NOS	11/11 (100.0%)	
Physical Impairment	11/11 (100.0%)	1.00***
Global Impairment Rating	8/11 (72.7%)	

* significant on 0.05 level ** significant on 0.01 level *** significant on 0.001 level

Four of the six CIA domains had a high percentage of agreement in the ratings at T1 and T2. These were Cognitive effects (90%), Work impairment (90%), Life NOS and Physical Impairment (both 100%). The kappa coefficients for Cognitive Effects and for Physical Impairment were "substantial" and "outstanding" respectively (Landis & Koch, 1977). Both were statistically significant. The domains Mood and Self-Perception and Interpersonal Impairment showed a moderate percentage of agreement (63.6% and 81.8%

respectively) with kappa values below a level of statistical significance. The CIA's main outcome, the global impairment rating, was rated identically at the two time points 72.7% of the time.

Summary of findings

- The global impairment rating on the CIA (i.e., whether or not a participant was classed as being clinically significantly impaired by their eating disorder symptoms) showed stability over time (and across raters) in 72.7% of the 11 cases interviewed on two occasions.
- Four of the six CIA life domains showed excellent test-retest reliability with the percentage agreement between ratings at the two time points ranging between 90% and 100%.
- The two life domains Mood and Self-Perception and Interpersonal Impairment showed moderate-to-good retest reliability, with the percentage agreement between ratings at the two time points being 63.6% and 81.8% respectively.
- The small sample size (N=11) needs to be taken into account when interpreting the findings of this study.
- This study was not a pure test of test-retest reliability, but a mix between the latter and inter-rater reliability (see Discussion, section 5.5 of this chapter).

5.3.3 Study Three: Construct and Discriminant Validity of the CIA

Design

The aim of this study was to test the construct and the discriminant validity of the CIA. As explained above, the CIA was not designed to provide an index of the severity of impairment secondary to eating disorder psychopathology, but rather to identify clinically

significant secondary impairment in a binary way. However, in order to test the validity of the instrument, two types of severity index were created from the CIA ratings.

1. The first was the total sum of ratings (on each of the six domains) which resulted in a total severity score ranging from 0 to 12.
2. The second was the creation of three degrees of severity of impairment. This was done on clinical grounds, in collaboration with Dr Helen Doll and Professor Christopher Fairburn. The three categories were:
    a. minimal or no impairment (defined as having no ratings of 2 and no more than two ratings of 1 across all the domains),
    b. moderate impairment (defined as having only one rating of 2 across the domains), and
    c. severe impairment (defined as having two or more ratings of 2 across the domains).

Table 5.7 gives an overview of the three outcomes of the CIA used in this study. They are

1. The global impairment rating (0 or 1)
2. The total severity score (range 0 to 12), and
3. The severity of impairment category (minimal, moderate, or severe).

Table 5.7 *Example of how the three outcomes of the CIA were created*

Patient X has received the following ratings in the six CIA domains:

Mood and Self-Perception	Cognitive Functioning	Interpersonal Functioning	Work Functioning	Life NOS	Physical Health
2	1	1	0	0	0

Since one of the domains was scored with a rating of '2', a rating of '1' is given on the *global impairment rating*, indicating that the secondary impairment is *clinically significant.*

The *impairment severity score* for this patient is *4* (2+1+1).

Since only one domain was scored with a rating of '2', the patient is placed in the *'moderate impairment' category.*

The test of the **construct validity** of the CIA involved examining whether the CIA was a valid measure of both the clinical significance and the severity of psychosocial and physical impairment secondary to eating disorder psychopathology. Ratings on the CIA (administered by a trained research assistant) were compared with impairment severity ratings made by one of two expert clinicians (Dr Zafra Cooper or Professor Christopher Fairburn), the ratings being of the same participants and covering the same period of time. The following three specific comparisons were made:

1. The clinicians' impairment severity ratings of those participants who were classed by the CIA to have clinically significant global impairment (referred to as the 'CS group' from this point onwards) were compared with the impairment severity ratings of those whose global impairment was not classed as clinically significant (referred to as the 'NS group').
2. Participants' total severity scores on the CIA (0 to 12) were compared with the clinicians' impairment severity ratings.
3. The clinicians' impairment severity ratings of participants showing a 'minimal', 'moderate' or 'severe' level of impairment (on the CIA) were compared.

With regard to these three comparisons it was predicted that:

1. The clinicians' impairment severity ratings of patients in the CS group would be statistically significantly higher than those of patients in the NS group.
2. There would be a statistically significant correlation between patients' total severity score on the CIA and the impairment severity ratings made by the clinicians.
3. The clinicians' impairment severity ratings of participants in the three CIA severity groups ('minimal', 'moderate' and 'severe') would differ significantly, with participants in the 'severe impairment' group having the highest and those in the 'minimal impairment" group having the lowest impairment ratings.

The **discriminant validity** of the CIA was tested in the following two ways.

1. The number of participants with an eating disorder of clinical severity (as judged by one of the two expert clinicians) in the CS group of the CIA was compared to that in the NS group.
2. The total CIA severity scores of patients with an eating disorder of clinical severity were compared with those of participants who were judged to no longer have a clinical eating disorder.

With regard to these two tests it was predicted that:

1. There would be statistically significantly more participants with a clinical eating disorder in the CIA CS group than in the NS group.
2. The CIA total severity scores of participants with a clinical eating disorder would be statistically significantly higher than those of participants without an eating disorder.

Participants

The participants were subsets of the 170 patients described in Chapter Three. There were four reasons for the use of subsets rather than the whole sample: first, the versions of

the CIA used in this study were only introduced after some patients had completed all the post-treatment assessments; second, only participants who presented with a 'non-trivial' degree of eating disturbance at the time of assessment were interviewed with the CIA (see section 5.3.5 above); third, participants suffering from a comorbid clinical depression at the time of assessment were excluded for the reason mentioned above; and fourth, the part of the study that involved testing the construct validity of the CIA was conducted at the Oxford site only, which reduced the number of potential participants to 105.

For the three tests of construct validity, the data of 66 of the 105 Oxford patients were used. These were the patients for whom both CIA data and a clinician's impairment rating were available at one or more assessment time points. For the two tests of discriminant validity, the data of 133 of the 170 patients described in Chapter Three were used. These were the patients for whom both CIA data and eating disorder caseness information were available at one or more assessment time points.

Measures

The CIA has been described in detail above (Section 5.2.5 of this chapter).

Clinicians' rating of secondary psychosocial and physical impairment

Two senior clinicians (Dr Zafra Cooper and Professor Christopher Fairburn) were asked to assess the extent to which the participants' disturbed eating habits and concerns about eating, shape or weight had been having an impact on their life over the previous 28 days. This was done as part of routine research assessments of each participant. These took place at the beginning and end of treatment, and at 20-, 40-, and 60-week and 2- and 3-year post treatment. Both senior clinicians were sent an e-mail by the candidate on the day an assessment was due reminding them to make three ratings: 1) the participant's DSM-IV eating disorder diagnosis over the past month; 2) and over the past three months; and 3) the

participant's degree of secondary psychosocial and physical impairment over the past month. Table 5.8 shows the content of this e-mail. The rating of the severity of secondary psychosocial and physical impairment followed the severity scoring system employed by the EDE. It consisted of a rating scale of 0 to 6 with higher ratings indicating a more severe level of secondary psychosocial impairment.

The clinicians were asked to make these ratings as soon as possible after their clinical assessment and to send them back to the candidate via e-mail.

Table 5.8 *Email sent to clinicians, asking them to rate the eating disorder case status and severity of secondary impairment of participants*

With regard to the forthcoming assessment of the above patient, please would you make the following THREE ratings:

RATING ONE - Eating disorder diagnosis based on DSM-IV criteria over past month
0 – Does not meet DSM criteria for an eating disorder
1 - Anorexia nervosa
2 - Bulimia nervosa
3 - BED
4 - Other EDNOS

RATING TWO - Eating disorder diagnosis based on DSM-IV criteria over last three months

RATING THREE - Impairment over past month
Severity of impairment due to the patient's eating habits, exercising, or feelings about eating, shape or weight OVER THE PAST MONTH.

Consider the following five domains:

- mood, view on self, guilt
- cognitive effects (preoccupation, concentration, forgetfulness)
- interpersonal/social effects (relationships with others, socialising, eating with others, sex)
- work (performance, time off)
- physical health

0 - No impairment
1
2 - Mild impairment
3 -
4 - Moderate impairment
5 -
6 - Severe impairment

Data analysis

Two different datasets were constructed for each validity sub-study (i.e., the tests of construct and discriminant validity). This was done in the following way: the first dataset was constructed by using all available and eligible CIA data on all participants (i.e., from all assessment times). This meant that each participant 'contributed' one or more pairs of ratings (i.e., CIA data and clinician's rating) to the dataset. This was done in order to increase the power of each study. The second dataset was constructed by using the last available CIA data[10] on each participant (i.e., the CIA data from the assessment time at which the participant last completed the CIA). This ensured that only one pair of ratings per participant was included in the analysis.

Construct validity

The first dataset for the three tests of construct validity consisted of 142 individual pairs of ratings from the 66 participants. The distribution of these pairs of ratings over the seven assessment points and the eating disorder case status of participants can be seen in Table 5.9. The second dataset consisted of 66 individual pairs of ratings (i.e., the last available CIA data for each of the participants, and the clinicians' severity ratings from the respective time point). Eighteen of the 66 participants (27.3%) had a clinical eating disorder at the time of assessment and 48 (72.7%) no longer had an eating disorder.

[10] The last available CIA data were chosen since doing so ensured that participants were on a spectrum with regard to the severity of their eating disorder features, since not all patients had completed treatment at the time of this study (whereas choosing the first available CIA data of each participant would have resulted in a too homogeneous dataset, since most participants were first assessed with the CIA before starting treatment).

Table 5.9 *First dataset for the construct validity study: Distribution of the 142 pairs of ratings over all assessment times and number of eating disorder cases and non-cases*

Assessment time	N (%)	Eating disorder cases	Non-cases
Before treatment	27 (19.0%)	27	
End of treatment	27 (19.0%)	4	23
20-week follow-up	29 (20.4%)	8	21
40-week follow-up	25 (17.6%)	6	19
60-week follow-up	21 (14.8%)	5	16
2-year follow-up	6 (4.2%)	1	5
3-year follow-up	7 (4.9%)	4	3
Total	**142 (100%)**	**55 (38.7%)**	**87 (61.3%)**

To test whether the participants who were identified by the CIA as having clinically significant secondary impairment (the CS group) received higher impairment severity ratings from the two clinicians than those who were classified as not significantly impaired (the NS group), a Mann-Whitney test was performed.

The relationship between total severity CIA score (range 0 to12) and the clinicians' impairment severity rating (range 0 to 6) was assessed using a two-tailed bivariate Spearman's correlation with alpha set at 0.01.

To test whether there was a statistically significant difference in the clinicians' ratings between the 'minimal', 'moderate' and 'severe' impairment groups as identified by the CIA, a Kruskal Wallis test and a one-way ANOVA including a Tukey B post hoc test were performed, based on an alpha coefficient of 0.01.

Discriminant validity

The first dataset for the two tests of discriminant validity consisted of 355 individual pairs of ratings (i.e., all available CIA data on 133 participants and clinicians' judgments about eating disorder caseness). The distribution of the pairs of ratings over the seven assessment time points and the eating disorder case status of participants is shown in Table 5.10. The second dataset consisted of 133 individual pairs of ratings. Fifty-eight of the 133

participants (43.6%) had an eating disorder of clinical severity at the time of assessment, and 75 (56.4%) did no longer have an eating disorder.

Table 5.10 *First dataset for the discriminant validity study: Distribution of the 355 pairs of ratings over all assessment times and number of eating disorder cases and non-cases*

Assessment time	**N (%)**	**Eating disorder cases**	**Non-cases**
Before treatment	83 (23.4%)	83	
End of treatment	60 (16.9%)	15	45
20-week follow-up	61 (17.2%)	21	40
40-week follow-up	67 (18.9%)	24	43
60-week follow-up	63 (17.7%)	21	42
2-year follow-up	14 (3.9%)	3	11
3-year follow-up	7 (2.0%)	4	3
Total	**355 (100%)**	**171 (48.2%)**	**184 (51.8%)**

To test whether there were more participants with an eating disorder of clinical severity in the CS group than in the NS group, a Chi-square test was performed. A Mann-Whitney test was performed to test whether the total CIA severity scores of participants with an eating disorder of clinical severity were higher than those of participants who were judged to no longer have a clinical eating disorder.

For ease of understanding the pairs of ratings in the first dataset of each study (in which participants' data are being used multiple times) are referred to as 'subjects' from this point onwards.

Results

Construct validity

First test of construct validity: Clinicians' impairment ratings in the CS and NS groups

First dataset (N=142)

The mean and standard deviation for the clinicians' impairment severity ratings in the CS and NS groups, and the result of their comparison are shown in Table 5.11. The table shows the p-value of the difference between the two ratings at each assessment point and for

all pairs of ratings. 'Subjects' who had been classed by the CIA to be clinically significantly impaired by their eating disorder features received statistically significantly higher impairment severity ratings from the clinicians than those who in the 'no significant impairment' group. This was true at every assessment point as well as for the total sample of measurements.

Second dataset (N=66)

Twenty-six of the 66 participants (39.4%) were classed as having clinically significant secondary psychosocial and physical impairment (CS group). The mean clinicians' rating for this group was 3.65 (SD=0.89). The 40 participants (60.6%) who showed no secondary impairment on the CIA or impairment at a not clinically significant level (NS group) received an average clinical rating of 1.50 (SD=1.13). This difference was statistically significant (Mann-Whitney 77.00; z=5.93; p<.001).

Table 5.11 *First test of construct validity: Comparison of clinicians' impairment ratings (mean and standard deviation) (range 0 to 6) for 'subjects' with and without clinically significant global impairment at each assessment point and in total (N=142)*

Assessment	N	CS[1]	N	NS[2]	Comparison of CS and NS (test statistic and p value)
Beginning of treatment	27	4.27 (0.78)			
End of treatment	8	3.50 (0.76)	19	1.53 (0.96)	z=3.64; p<0.001***
20-week follow-up	11	3.09 (1.04)	18	1.44 (0.98)	z=3.35, p=0.001***
40-week follow-up	7	3.43 (0.98)	18	1.44 (1.15)	z=3.22; p=0.001***
60-week follow-up	6	3.17 (0.75)	15	1.27 (1.16)	z=2.99; p=0.003**
2- and 3year follow-up	8	3.75 (0.89)	5	2.20 (0.84)	z=2.44; p=.015*
Total	**67**	**3.73 (0.95)**	**75**	**1.51 (1.08)**	**z=8.92; p<0.001*** **

* significant on 0.05 level
** significant on 0.01 level
*** significant on 0.001 level

[1] Subjects showing clinically significant global impairment
[2] Subjects without clinically significant global impairment

Second test of construct validity: Relationship between the CIA total severity score and the clinicians' impairment ratings

First dataset (N=142)

The total CIA severity scores and clinicians' ratings, and the correlation coefficients are shown in Table 5.12. The table shows the correlation coefficient for each time point and for all 'subjects' (total). There was a statistically significant positive correlation between the two measures at six of the seven assessment points. This was also true for the total sample of 'subjects'. Thus, a higher total CIA score was associated with a higher rating of secondary psychosocial and physical impairment as judged by the clinician.

Table 5.12 *Second test of construct validity: Spearman's correlation for CIA total severity score and clinicians' rating*

Assessment time	N	CIA total score M (SD)	Clinician's rating M (SD)	r (p)
Beginning of treatment	27	5.49 (1.91)	4.26 (0.76)	.87 (<.001)***
End of treatment	27	1.36 (1.77)	2.11 (1.28)	.70 (<.001)***
20-week follow-up	29	2.17 (2.45)	2.07 (1.28)	.74 (<.001)***
40-week follow-up	25	1.78 (2.01)	2.00 (1.41)	.72 (<.001)***
60-week follow-up	21	1.44 (2.02)	1.81 (1.37)	.71 (<.001)***
2-year follow-up	6	3.67 (2.94)	2.50 (1.05)	.67 (.144)
3-year follow-up	7	3.06 (1.96)	3.71 (0.95)	.77 (.042)*
Total	**142**	**2.58 (2.55)**	**2.54 (1.51)**	**.83 (<.001)*** **

* significant on 0.05 level
** significant on 0.01 level
*** significant on 0.001 level

Second dataset (N=66)

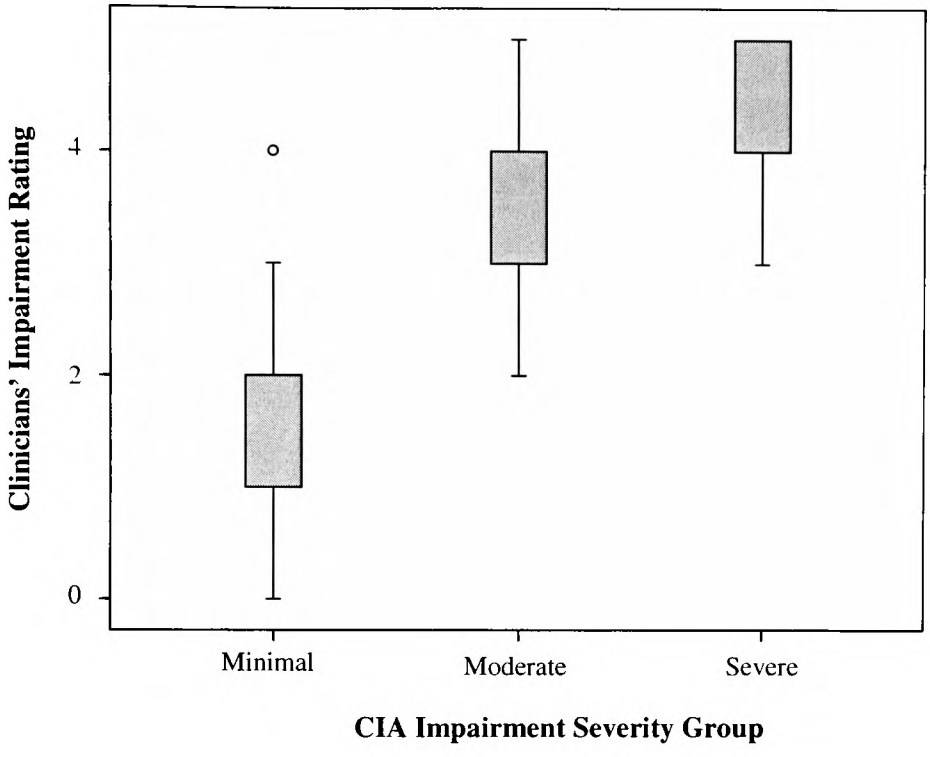
The mean total CIA severity score for the 66 participants was 2.02 (SD=2.25). The average impairment rating these participants received from the expert clinicians was 2.35 (SD=1.48). The Spearman's correlation between these two measures was statistically significant (r=.80, p<.001).

Third test of construct validity: Relationship between the three CIA severity ratings and the clinicians' impairment ratings

First dataset (N=142)

Of the 142 'subjects', 76 (53.5%) were in the 'minimal' impairment group (i.e., they received no more than two ratings of 1 in any of the six CIA domains); 35 (24.6%) were in the 'moderate' impairment group (i.e., they received no more than one rating of 2 in any of the five domains), and 31 (21.8%) fell into the 'severe' impairment group (i.e., they received two or more ratings of 2 across the five CIA domains). Figure 5.2 shows a boxplot diagram of the distribution of clinicians' impairment ratings per CIA severity rating group. It can be seen that 'subjects' who were in the 'severe' CIA impairment group had the highest mean clinical impairment rating (4.26; SD=0.77). The 'subjects' in the 'moderate' impairment group had a mean clinicians' impairment rating of 3.26 (SD=0.58), while those in the 'minimal' impairment group had a mean clinical rating of 1.51 (SD=1.08). The Kruskal Wallis test revealed that there was a statistically significant difference in the clinical ratings between the three CIA severity groups ($F(2.0)=86.46$, $p<0.001$). The one-way ANOVA confirmed this finding ($F(2.0)=101.69$, $p<0.001$). The Tukey B post-hoc test revealed that the clinicians' ratings in each CIA severity group differed from those in every other severity group at a $p=0.05$ level.

Figure 5.2 *Third test of construct validity: Clinicians' impairment ratings (0-6) in the three CIA impairment severity groups*



Second dataset (N=66)

The same difference in clinicians' ratings between the three CIA severity groups was found when only one set of data per participant was included in the analysis. The 'minimal impairment' group (N=40; 60.6%) had the lowest mean clinicians' rating (1.50; SD=1.13), the 'moderate impairment' group (N=17; 25.8%) had a mean rating of 3.41 (SD=0.80), and the 'severe impairment' group (N=9; 13.6%) had the highest mean clinicians' rating (4.11; SD=0.93). Both, the Kruskal Wallis and the one-way ANOVA showed that these differences were statistically significant (Kruskal Wallis: $F(2)=36.12$, $p<0.001$; ANOVA: $F(2.0)=35.718$, $p<0.001$). The Tukey post-hoc test revealed that the 'minimal' impairment group had significantly lower ratings than the two other severity groups. The clinicians' ratings in the moderate and severe impairment groups did not differ significantly.

Discriminant validity

First test of discriminant validity: Eating disorder cases in the CIA CS and NS groups

First dataset (N=355)

Of the 355 'subjects', 195 (54.9%) were classed by the CIA as clinically significantly impaired by their eating disorder symptoms (CS group), whereas 160 (45.1%) did not have significant global impairment (NS group). The great majority of 'subjects' in the CS group (N=159; 81.5%) were judged (by the two expert clinicians) to have an eating disorder of clinical severity, compared to only 7.5% (N=12) in the NS group. A Chi-square test revealed that this difference was statistically significant (χ^2=190.01, df=1, p<0.001). This was also true for every assessment time point, with the exception of the beginning of treatment (when every participant had an eating disorder of clinical severity) and the 2- and 3-year follow-up (where only 21 pairs of ratings were available).

Second dataset (N=133)

Sixty-seven of the 133 participants (50.4%) were classed as belonging to the CS group of the CIA whereas 66 (49.6%) fell into the NS group. There were significantly more participants with a clinical eating disorder in the CS group compared to the NS group (79.1% vs. 7.6%) (χ^2=66.30, df=1, p<0.001).

Second test of discriminant validity: Relationship between eating disorder caseness and the CIA total severity score

First dataset (N=355)

The CIA total severity scores of subjects with and without a clinical eating disorder, and their comparison, are shown in Table 5.13. The table shows the p-value of the difference between the two scores at each assessment point and for all pairs of ratings. 'Subjects' judged to have an eating disorder of clinical severity had statistically significantly

higher total CIA severity scores than those without an eating disorder. This was true at every assessment point (with the exception of the two and three-year follow-up) as well as for the total sample of subjects.

Table 5.13 *Second test of discriminant validity: Comparison of CIA total severity scores (mean and standard deviation) (range 0 to 12) for participants with and without a clinical eating disorder*

	N	ED cases	N	Non-cases	Comparison of the two groups (test statistic and p value)
Assessment time					
Before treatment	83	4.95 (1.60)			
End of treatment	15	5.10 (2.15)	45	0.82 (1.40)	z=5.56, p<0.001***
20-week follow-up	21	4.12 (2.42)	40	1.18 (1.65)	z=4.53, p<0.001***
40-week follow-up	24	4.83 (1.98)	43	1.02 (1.59)	z=6.07, p<0.001***
60-week follow-up	21	4.80 (1.48)	42	0.89 (1.45)	z=6.09, p<0.001***
2- and 3year follow-up	7	4.00 (3.21)	14	2.82 (1.63)	z=1.02, p=0.31
Total	**171**	**4.79 (1.89)**	**184**	**1.11 (1.60)**	**z=13.94, p<0.001*** **

*** significant on 0.001 level

Second dataset (N=133)

Fifty-eight of the 133 participants (43.6%) were judged to have a clinical eating disorder. The mean CIA total severity score for this group was 4.61 (SD=1.82). The 75 participants (56.4%) who were judged no longer to have an eating disorder had an average CIA total severity score of 1.01 (SD=1.37). This difference was statistically significant (Mann-Whitney 277.00; z=8.80; p<.001).

Summary of findings

Construct validity of the CIA:

- Participants who were judged to be clinically significantly impaired by their eating disorder symptoms on the CIA received statistically significantly higher impairment ratings by an expert clinician than those who did not show significant impairment on the CIA.

- There was a strong positive correlation between participants' total severity scores on the CIA and the impairment ratings made by an expert clinician.
- There was a statistically significant difference in the clinicians' impairment ratings between the three CIA severity groups. Participants in the 'severe impairment' group received the highest impairment ratings by a clinician, followed by the participants in the 'moderate' and the 'minimal' impairment groups.

Discriminant validity of the CIA:

- About eighty percent of the participants who were judged to be clinically significantly impaired on the CIA had an eating disorder of clinical severity. This was true for only 7.5% of the participants who did not show any significant impairment on the CIA.
- The total severity scores on the CIA were statistically significantly higher in participants with an eating disorder of clinical severity than in those without one.

All of the above findings were independent of whether all available data were entered into the analyses (first dataset) or only one set of data per participant (second dataset).

These findings suggest that the CIA is a valid measure of the clinical significance and the severity of secondary functional impairment due to eating disorder features and that it successfully discriminates between people with and without an eating disorder.

5.4 Identifying those eating disorder features that are associated with clinically significant secondary functional impairment

5.4.1 Introduction

In this section of the chapter a study is described that utilised the newly developed Clinical Impairment Assessment (CIA) interview to identify the type and level of eating

disorder psychopathology that is associated with clinically significant impairment of functioning, thereby generating a provisional operational definition of an eating disorder.

5.4.2 Design

The study had two aims. The first was to identify which specific eating disorder features are statistically significantly associated with a clinically significant level of secondary functional impairment. The second was to identify the levels of severity of these features that best predict being clinically significantly impaired. The design involved administering the EDE and the CIA to a group of subjects with varying degrees of eating disorder psychopathology (i.e., exhibiting the complete spectrum of severity of eating disorder features). The eating disorder features (as assessed using the EDE) of participants who were classed on the CIA as clinically significantly impaired were compared with those of participants who were not classed as clinically significantly impaired.

5.4.3 Participants

The sample comprised 127 of the 170 patients with a clinical eating disorder described in Chapter Three. These were the patients who a) at the time when the analysis for the current study was conducted (May 2006) had been administered the CIA (version 2.0, 2.1 or 3.0) at one or more assessment times, b) did show "non-trivial" levels of eating disorder psychopathology, and c) did not suffer from a comorbid clinical depression. As described above, each was assessed with the EDE and the CIA as part of their research assessment at one or more time points before and after receiving treatment for their eating disorder (viz., beginning and end of treatment, and at 20-, 40-, and 60-week and two- and three-year post-treatment follow-up). They were also assessed by one of three expert clinicians (see Section 5.3 above) at each of these times who determined whether the person had an eating disorder of clinical severity.

It was decided to construct the dataset for this study by identifying all available pairs of data (EDE data and CIA ratings) on the 127 participants, rather than using just one set of data per participant. This was done in order to increase the power of the study. It resulted in a total of 334 sets of data. Table 5.14 shows the distribution of these datasets over the various assessment times and the number of datasets with and without clinically significant impairment (i.e., the number of datasets in the CS and the NS group of the CIA). It also shows participants' eating disorder case status.

Of the 334 sets of data, 192 (57.5%) contained a CIA global impairment rating of 1 (i.e., a clinically significant level of secondary impairment) and 142 (42.5%) a CIA global impairment rating of 0 (i.e., mild or no secondary impairment). One hundred and sixty-two of the 334 pairs of ratings (48.5%) were acquired at time points when the participants had a clinical eating disorder and 172 (51.5%) at times when participants did not have an eating disorder of clinical severity.

For ease of understanding the 334 'pairs of ratings' are referred to as 'subjects' from this point onwards.

Table 5.14 *Distribution of the 334 pairs of rating over the five assessment points and number of pairs of ratings ('subjects') in the CS[1] and NS[2] groups and their eating disorder case status*

	CS[1] (N=192)	NS[2] (N=142)	Total (N=334)
Assessment time (N, %)			
Before treatment	83		83 (24.9%)
End of treatment	25	35	60 (18.0%)
20-week follow-up	29	32	61 (18.3%)
40-week follow-up	29	38	67 (20.1%)
60-week follow-up	26	37	63 (18.9%)
Total	**192 (57.5%)**	**142 (42.5%)**	**334 (100%)**
Presence of a clinical eating disorder[3]	157 (47.0%)	5 (1.5%)	162 (48.5%)

[1] Subjects showing clinically significant global impairment
[2] Subjects without clinically significant global impairment
[3] As assessed by an expert eating disorder clinician

5.4.4 Measures

The two measures in this study, the EDE and the CIA, have been described earlier (EDE in section 3.2.2 of Chapter Three; CIA in section 5.2.5 of this chapter). Version 15.0 of the EDE was used in this study. It is shown in Appendix 3.1. Versions 2.0, 2.1 and 3.0 of the CIA were used in this study. Version 3.0 is shown in Appendix 5.1.

The EDE was chosen as the measure of participants' eating disorder features since it is widely viewed as the gold standard for assessing the full range of eating disorder psychopathology (Garner, 1995). The time frame chosen for the assessment of eating disorder psychopathology and secondary functional impairment (as assessed with the CIA) was the last 28 days. This was done since it is the standard EDE time frame and since information recalled from over this time period is likely to be more valid than that recalled from two or three months ago.

5.4.5 Data analysis

5.4.5.1 Pre-selection of suitable EDE items

Of the 46 individual EDE items (EDE version 15.0; see Appendix 3.1), 30 were selected for a preliminary analysis to determine whether there were specific eating disorder features that were individually associated with being clinically significantly impaired. The remaining 16 EDE items were excluded for conceptual reasons (see Appendix 5.3). Two composite variables were added. These were: the total sum of bulimic episodes ("all bulimic episodes" (ABEs); the sum of OBEs and SBEs) and episodes of 'purging' (the sum of episodes of self-induced vomiting and laxative misuse).

The following 32 variables were entered into the preliminary analysis:

- 5 Restraint subscale items[11]

  Restraint over eating

[11] The 'total' ratings are used for these items (i.e., for shape/weight and/or control reasons).

Avoidance of eating

Empty stomach

Food avoidance

Dietary rules

- 7 Shape Concern subscale Items

  Dissatisfaction with shape

  Importance of shape[12]

  Fear of weight gain

  Discomfort seeing body

  Discomfort about exposure

  Feelings of fatness

  Flat stomach

- 4 Weight Concern subscale items

  Dissatisfaction with weight

  Desire to lose weight

  Reaction to prescribed weighing

  Importance of weight[12]

- Objective bulimic episodes
- All bulimic episodes
- Episodes of objective overeating
- Episodes of compensatory self-induced vomiting
- Episodes of purging
- Episodes of laxative misuse
- Average number of laxatives taken on each occasion
- Episodes of diuretic misuse
- Average number of diuretics taken on each occasion
- Driven exercising (number of days)
- Weighing
- Importance of strict control over eating

[12] The 'unadjusted' ratings are used for these items.

- Size discomfort
- Regional fatness
- Vigilance about shape
- Body mass index

To identify whether there were specific eating disorder features that predicted being clinically significantly impaired, the responses on these 32 items of 'subjects' with clinically significant impairment (N=192) were compared with the responses of those without (N=142). Chi-square statistics (for dichotomous items) and the Mann-Whitney test (for continuous items) were used to determine the statistical significance of any apparent association between individual items and impairment status. In addition it was examined whether the time point of assessment had an impact on the association between items and impairment status by comparing subjects' responses on the 32 items at different assessment points.

The EDE items that were found to be significantly associated with a clinically significant level of impairment were then subjected to an exploratory logistic regression analysis (performed using SPSS, version 14.0), using forward and backward stepwise selection of covariates. This analysis was undertaken in order to detect the EDE items with the strongest independent effects for the subsequent more detailed analysis.

5.4.5.2 Detailed analysis

The EDE items that had been found to be significantly associated with impairment status in the preliminary analysis were subjected to three multiple variable analyses: logistic regression, linear discriminant function analysis, and signal detection analysis. The former two analyses were performed within SPSS (version 14.0), and both used the forward stepwise selection of covariates. The signal detection analysis was performed with ROC4.19

(Mental Illness Research Educational and Clinical Centre, Stanford, Conn.; ROC4.19 is available for download at http://mirecc.stanford.edu).

Logistic regression

This analysis was performed to determine which eating disorder features were independently predictive of being clinically significantly impaired.

Linear discriminant function analysis

This analysis was performed to identify, through the creation of a discriminant function, which eating disorder features best discriminated clinically significantly impaired subjects from those who were not impaired. Standardised discriminant function coefficients and variable loadings are presented.

Signal detection analysis

This analysis was used to determine, through creation of a decision tree, the most sensitive and specific algorithm to identify subjects who are clinically significantly impaired. Agras and colleagues (Agras, Scott, Halmi, Mitchell, Wilson & Kraemer, 2000) summarise the process of signal detection analysis in the following way:

> *"In a first step, signal detection considers each possible predictor (including a range of different cutoff points for any ordinal predictor). For each, it computes the sensitivity and the specificity of that "test" against the outcome. Using the selected weighting of the relative clinical importance of false positives and false negatives, it finds the optimal predictor (and optimal cut-off point for an ordinal predictor). This is then used to split the initial population into two subsets, the one positive on the first "test" and one that is negative. The process is repeated on each of these two subsets. (...) The process is then repeated on each of the resulting four subsets, then on the resulting eight, etc., ultimately creating a decision tree. The process stops when there are no more "tests", when the sample size in some subset is too small, or when the optimal test does not achieve some preset criterion (...)."*
>
> (Agras et al., 2000, p.1304)

The analysis was performed with a 50:50 emphasis on sensitivity versus specificity as there was no a priori reason to emphasise either sensitivity or specificity.

5.4.5.3 Determination of cut-off points

A final logistic regression analysis was performed to determine which severity levels on the EDE items that had been found to be associated with impairment status (by at least two of the three preceding analyses) were most strongly associated with a clinically significant level of impairment.

5.4.6 Results

One-hundred and ninety-two of the 334 'subjects' (57.5%) were classed as being clinically significantly impaired on the CIA (a global impairment rating of 1), while 142 (42.5%) were judged not to have secondary impairment at a clinical level.

5.4.6.1 Pre-selection of suitable EDE items

Table 5.15 shows the comparisons between 'subjects' with clinically significantly impairment (CS group) and those without (NS group) on the 32 EDE items that were entered into the preliminary analysis. Subjects in the CS group scored higher than those in the NS group on all EDE items, with the exception of body mass index. Chi square values are presented for categorical EDE items and Mann-Whitney values for continuous EDE items. Chi square tests were calculated for unrecoded items (0-6) and for a rating of 0 to 3 versus a rating of 4 to 6. P values are presented for the whole sample of subjects and for subjects at two different assessment time frames, the first one being the end-of treatment or 20-week post-treatment follow-up, and the second at 40- or 60-week post-treatment follow-up. The beginning of treatment was not examined as a separate assessment time point, since all subjects presented with clinically significant impairment at this assessment time. Items in bold did not discriminate between the two groups.

As can be seen in Table 5.15, 27 of the 32 EDE items were significantly associated with impairment status at a p-level of 0.05, when the whole sample of subjects was entered

into the analysis. The subjects in the CS group scored statistically significantly higher on these items than those in the NS group. The following five items did not show a statistically significant association with impairment status: Avoidance of eating, Objective overeating episodes, Average number of laxatives taken, Episodes of diuretic misuse and Average number of diuretics taken. All of these features only occurred in a minority of subjects. They were excluded from further analysis.

Two further EDE items were dropped from further consideration. The item 'Empty stomach' also only appeared in a small minority of subjects, and the item 'Size discomfort' contained too much missing data.

	Comparison of participants in CS[1] and NS[2] groups **Test statistic and p value**			
EDE items	**All 'subjects'** **(N=334; CS=192, NS=142)**		**End-of-tr. and 20-wk f-u** **(N=121; CS=54, NS=67)**	**40- and 60-wk f-u** **(N=130; CS=55, NS=75)**
	Un-recoded items	***Rating ≤3 vs >3*** ***(for categorical items)***	***Rating ≤3 vs >3*** ***(for categorical items)***	***Rating ≤3 vs >3*** ***(for categorical items)***
Restraint	139.12; 7; p<0.001	χ^2=110.31; df=1; p<0.001	22.15; 1; p<0.001	42.78; 1; p<0.001
Avoidance of eating	-	**χ^2=2.20; df=1; p=0.14**	-	-
Empty stomach	-	χ^2=25.86; df=1; p<0.001	1.91; 1; p=0.17	7.36; 1; p=.007
Food avoidance	118.49; 6; p<0.001	χ^2=84.10; df=1; p<0.001	15.40; 1; p<0.001	31.83; 1; p<0.001
Dietary rules	-	χ^2=122.49; df=1; p<0.001	32.58; 1; p<0.001	39.82; 1; p<0.001
Dissatisfaction with shape	96.84; 6; p<0.001	χ^2=89.15; df=1; p<0.001	40.81; 1; p<0.001	22.08; 1; p<0.001
Importance of shape	87.92; 6; p<0.001	χ^2=61.00; df=1; p<0.001	13.26; 1; p<0.001	14.05; 1; p<0.001
Fear of weight gain	98.20; 7; p<0.001	χ^2=70.67; df=1; p<0.001	17.01; 1; p<0.001	20.17; 1; p<0.001
Discomfort seeing body	70.88; 6; p<0.001	$\chi^2$46.77; 1; p<0.001	12.77; 1; p<0.001	17.75; 1; p<0.001
Discomfort about exposure	65.04; 6; p<0.001	$\chi^2$49.08; 1; p<0.001	14.05; 1; p<0.001	21.39; 1; p<0.001
Feelings of fatness	69.82; 7; p<0.001	$\chi^2$54.82; 1; p<0.001	10.03; 1; p=.002	27.00; 1; p<0.001
Flat stomach	67.86; 6; p<0.001	$\chi^2$54.86; 1; p<0.001	6.67; 1; p=.01	17.94; 1; p<0.001
Dissatisfaction with weight	-	$\chi^2$67.18; 1; p<0.001	22.39; 1; p<0.001	21.61; 1; p<0.001
Desire to lose weight	63.79; 6; p<0.001	$\chi^2$51.88; 1; p<0.001	10.65; 1; p=.001	17.86; 1; p<0.001
Reaction to prescribed weighing	61.48; 6; p<0.001	$\chi^2$49.76; 1; p<0.001	5.63; 1; p=.018	15.24; 1; p<0.001
Importance of weight	103.29; 7; p<0.001	$\chi^2$78.42; 1; p<0.001	18.27; 1; p<0.001	24.71; 1; p<0.001
Objective bulimic episodes	z=7.53; p<0.001	n/a	Z=2.57; p=0.01	z=5.38; p<0.001
ABE	z=10.04; p<0.001	n/a	Z=4.41; p=<0.001	z=5.58; p<0.001
OOE	**z=0.80; p=0.42**	**n/a**	**Z=0.77; p=0.44**	**z=0.08; p=0.93**
SIV episodes	z=7.20; p<0.001	n/a	Z=3.27; p=0.001	z=3.73; p<0.001
Laxative episodes	z=3.60; p<0.001	n/a	Z=1.11; p=0.27	z=2.34; p=0.02
Average no laxs taken	**z=1.29; p=0.20**	**n/a**	-	**z=1.33; p=0.19**
Diuretic episodes	**z-1.16; p=0.25**	**n/a**	**Z=0.90; p=0.37**	**z=0.00; p=1.0**
Average no diur taken	**-**	**n/a**	**-**	**-**
Driven exercise (days)	z=5.03; p<0.001	n/a	Z=3.19; p=0.001	z=3.34; p=0.001
Weighing	z=5.16; p<0.001	n/a	Z=3.55; p<0.001	z=2.61; p=0.009
Importance of strict control	-	$\chi^2$76.86; 1; p<0.001	19.92; 1; p<0.001	17.53; 1; p<0.001
Size discomfort	-	χ^2 14.39; 1; p<0.001	5.08; 1; p=0.02	-

Regional fatness	47.83; 7; p<0.001	$\chi^2$38.52; 1; p<0.001	9.12; 1; p=.003	10.74; 1; p=.001
Vigilance about shape	-	$\chi^2$26.53; 1; p<0.001	7.11; 1; p=.008	6.70; 1; p=.01
Body mass index	z=3.33; p=0.001	n/a	Z=1.59; p=0.11	z=2.28; p=0.23

[1] Subjects with clinically significant global impairment on the CIA
[2] Subjects without clinically significant global impairment on the CIA

Table 5.15 *Comparison of subjects with clinically significant impairment and those without on 32 EDE items (for total sample, end-of-treatment and 20-week follow-up, and 40- and 60-week follow-up)* [NB: items in bold did not discriminate between the two groups]

The remaining 25 EDE items had a statistically significant association with impairment status that was consistent over assessment times, the only exception being body mass index (BMI). BMI was significantly associated with being impaired only when the whole sample of subjects was considered. Thus just these 25 items were entered into the exploratory logistic regression analysis.

Exploratory logistic regression analysis

Two hundred and seventy-one 'subjects' (81.1%) were included in this analysis. The remaining 63 had missing values on one or more of the EDE items and were therefore excluded. The 25 selected EDE items were entered using Wald's forward and backward stepwise selection of predictor variables. This means that, at each step, the EDE item (predictor variable) with the strongest association with impairment status enters the model. This process is repeated until there are no remaining predictor variables with a strong association ($p<0.05$) with impairment status.

Five steps were performed in total. The Omnibus Tests of Model Coefficients showed that step five had a good 'goodness of fit' ($\chi^2=10.24$; df=1; p=0.001). The Cox & Snell R Square and the Nagelkerke R Square values indicated that between 56.2% and 75.1% of the variation in impairment status were explained by the five EDE items selected in step five. The five EDE items, selected by logistic regression analysis to be most strongly associated with impairment status, were:

- Dietary rules
- All bulimic episodes (ABEs)
- Episodes of purging
- Dissatisfaction with shape
- Body mass index

The model correctly classified 88.2% of the 271 subjects. The sensitivity and specificity of the model were both high. The sensitivity (i.e. the percentage of subjects that showed clinically significant impairment on the CIA that were accurately identified by the model) was 87.4%. Its specificity (i.e., percentage of subjects who were not impaired on the CIA correctly predicted not to be impaired by the model) was 89.1%.

In addition to these five EDE items, it was decided to retain six further EDE items for more detailed analysis. This was done since these additional six items showed a strong association with impairment status ($p<0.001$), after the model had adjusted for the above five variables. These six EDE items were:

- Objective bulimic episodes
- Importance of weight
- Importance of shape
- Importance of strict control over eating
- Dissatisfaction with weight
- Desire to lose weight

For conceptual reasons it was decided to create the following two composite variables: "Dissatisfaction with weight and/or shape" and "Importance of weight and/or shape". The two variables were created by using the highest rating of the two pairs of ratings (i.e., either the shape or weight rating for each variable).

This resulted in nine EDE items being entered into the detailed analysis. Their distribution among 'subjects' with and without clinically significant impairment is shown in Table 5.16.

Table 5.16 *Preliminary analysis: Distribution of nine EDE items selected for the detailed analysis amongst 'subjects' with and without clinically significant functional impairment*

EDE items	CS group[1] N=192 n (%)	NS group[2] N=142 n (%)	Comparison between CS and NS group
Dietary rules			χ^2=141.65; df=6; p<0.001
0	45 (24.1)	123 (86.6)	
1	5 (2.7)	7 (4.9)	
2	9 (4.8)	2 (1.4)	
3	5 (2.7)	3 (2.1)	
4	17 (9.1)	1 (0.7)	
5	11 (5.9)	0 (0)	
6	95 (50.8)	6 (4.2)	
OBE[3] (mean, SD)	8.41 (13.89)	0.56 (2.07)	z=7.53; p<0.001
All binge episodes (ABE) (mean, SD)	20.68 (36.57)	2.32 (14.33)	z=10.04; p<0.001
Desire to lose weight			χ^2=63.79; df=6; p<0.001
0	41 (21.6)	80 (56.7)	
1	13 (6.8)	16 (11.3)	
2	17 (8.9)	10 (7.1)	
3	11 (5.8)	11 (7.8)	
4	10 (5.3)	3 (2.1)	
5	10 (5.3)	0 (0)	
6	88 (46.3)	21 (14.9)	
Importance of strict control			χ^2=103.71; df=6; p<0.001
0	55 (29.6)	115 (82.1)	
1	1 (0.5)	1 (0.7)	
2	20 (10.6)	14 (10.0)	
3	11 (5.9)	1 (0.7)	
4	31 (16.7)	7 (5.0)	
5	37 (19.9)	2 (1.4)	
6	31 (16.7)	0 (0)	
BMI[4] (mean, SD)	22.58 (4.55)	23.80 (4.17)	z=3.23; p=0.001
Purging episodes (mean, SD)	13.04 (26.63)	0.70 (2.15)	z=7.54; p<0.001
Dissatisfaction with shape and/or weight			χ^2=110.61; df=6; p<0.001
0	14 (7.3)	33 (23.2)	
1	0 (0)	7 (4.9)	
2	6 (3.1)	35 (24.6)	
3	25 (13.0)	37 (26.1)	
4	81 (42.2)	20 (14.1)	
5	46 (24.0)	7 (4.9)	
6	20 (10.4)	3 (2.1)	
Importance of weight and/or shape			χ^2=109.03; df=6; p<0.001
0	3 (1.6)	9 (6.3)	
1	0 (0)	3 (2.1)	
2	5 (2.6)	37 (26.1)	
3	11 (5.8)	27 (19.0)	
4	40 (20.9)	43 (30.3)	
5	56 (29.3)	10 (7.0)	
6	76 (39.8)	13 (9.2)	

[1] Subjects with clinically significant global impairment on the CIA
[2] Subjects without clinically significant global impairment on the CIA
[3] Objective bulimic episodes
[4] Body mass index

5.4.6.2 Detailed analysis

Logistic regression analysis

Three-hundred and fourteen 'subjects' (94%) were included in this analysis. The remaining 20 had missing values on some of the selected EDE items and were therefore excluded. The nine selected EDE items (see Table 5.16) were entered into a second logistic regression analysis using Wald's forward stepwise selection of predictor variables. Five steps were performed in total. The Omnibus Tests of Model Coefficients showed that step five had a good 'goodness of fit' (χ^2=14.76; df=6; p=0.02). The Cox & Snell R Square and the Nagelkerke R Square values indicated that between 56.2% and 75.4% of the variation in impairment status were explained by the five EDE items selected in step five. These EDE items are shown in Table 5.17 below. The model correctly classified 89.8% of the 314 subjects. The sensitivity and specificity of the model were both high (89.8% and 89.9% respectively).

Table 5.17 shows the five variables that were most strongly associated with the presence of clinically significant impairment and their contribution to the model. The variables are shown in the order in which they entered the model. In addition, severity levels of the categorical variables and their importance with regard to the predictive ability of the model are shown. For example, it can be seen that a rating of 4 on the Dietary Rules item of the EDE was significantly associated with impairment (p=.025). The Exp(B) value, the odds ratio, indicates that a person was 15 times more likely to be impaired if they had a rating of 4 on the Dietary Rules item than if they did not show this eating disorder feature. The 95% confidence interval shows the region within which there is a 95% probability that the true

value lies. Wide confidence intervals are likely to reflect small sample sizes. The positive B value indicates that the presence of this variable results in an increased (rather than decreased) probability of being clinically significantly impaired.

Table 5.17 *Detailed analysis: EDE items selected by logistic regression analysis to be most strongly associated with the presence of clinically significant impairment and the strength of association for each level of severity (ratings 1-6)*

						95% C.I. for Exp(B)	
EDE items	**B**	**Wald**	**Df**	**Sig.**	**Exp(B)**	**Lower**	**Upper**
Dietary rules		21.224	6	.002			
1	.790	1.031	1	.310	2.204	.480	10.131
2	1.561	3.027	1	.082	4.763	.821	27.634
3	.712	.527	1	.468	2.037	.298	13.908
4	2.710	5.056	1	.025	15.029	1.416	159.530
5	20.566	.000	1	.999		.000	
6	2.519	17.361	1	.000	12.411	3.796	40.582
OBE[1]	.204	9.192	1	.002	1.226	1.075	1.398
Importance of strict control		6.720	6	.348			
1	-2.167	.039	1	.843	.115	.000	
2	.543	.974	1	.324	1.722	.585	5.064
3	2.605	4.488	1	.034	13.538	1.215	150.813
4	.923	1.883	1	.170	2.518	.673	9.414
5	1.036	1.038	1	.308	2.818	.384	20.665
6	19.322	.000	1	.997		.000	
Purging	.142	4.033	1	.045	1.153	1.003	1.324
Dissatisfaction with shape/weight		17.338	6	.008			
1	-19.17	.000	1	.999	.000	.000	
2	-1.171	2.222	1	.136	.310	.067	1.446
3	-.498	.594	1	.441	.608	.171	2.158
4	.953	2.472	1	.116	2.593	.791	8.503
5	1.459	3.807	1	.051	4.303	.993	18.642
6	1.492	2.556	1	.110	4.447	.714	27.698

[1] Objective bulimic episodes

Discriminant function analysis

As in the logistic regression analysis, 314 of the 334 'subjects' (94%) were included in this analysis. The nine selected EDE items (see Table 5.16) were entered into the analysis

using a forward stepwise selection of variables. In stepwise discriminant function analysis, a model of discrimination is built step-by-step. Specifically, at each step all variables are reviewed and evaluated to determine which one will contribute most to the discrimination between groups. That variable will then be included in the model, and then the process starts again.

Five steps were performed by the analysis. The resulting canonical discriminant function (containing five variables) had an eigenvalue of 1.165 and a Wilk's Lambda of 0.462. The eigenvalue was above 1, which meant that a large proportion of the variance in the dependent variable (here impairment status) was explained by the identified function. Wilk's Lambda tests the significance of the eigenvalue for the discriminant function. In this case it was statistically significant (χ^2=239.10; df=5; p<0.001).

Table 5.18 shows the five variables that were entered into the model. Also shown in the table are standardised discriminant function coefficients for each variable. The latter indicate the relative importance of each variable in predicting impairment status. The larger the standardised coefficient, the greater is the contribution of the respective variable to the discrimination between impairment and non-impairment status. Furthermore, the correlations of each variable with the discriminant function (also referred to as variable loading) can be seen in the table. The discriminant function correctly classified 84.1% of the 334 subjects. The sensitivity and specificity of the model were 80.2% and 89.4% respectively.

Discriminant function analysis makes a number of assumptions about the data that should not be violated, one of which is the homogeneity of variances and covariances of the variables entered. Box's test of equality of covariance matrices tests this assumption. The test for the current dataset was significant (Box's M=250.89; p<0.001), which meant that the

two groups (impairment and non-impairment) differed in their covariance matrices, violating an assumption of discriminant function analysis. However, it has been shown that discriminant function analysis is robust even when the homogeneity of variances assumption is not met, provided the data do not contain extreme outliers.

Table 5.18 *Detailed analysis: EDE items identified by discriminant function analysis to be most predictive of being clinically significantly impaired and their variable loading and standardised coefficients*

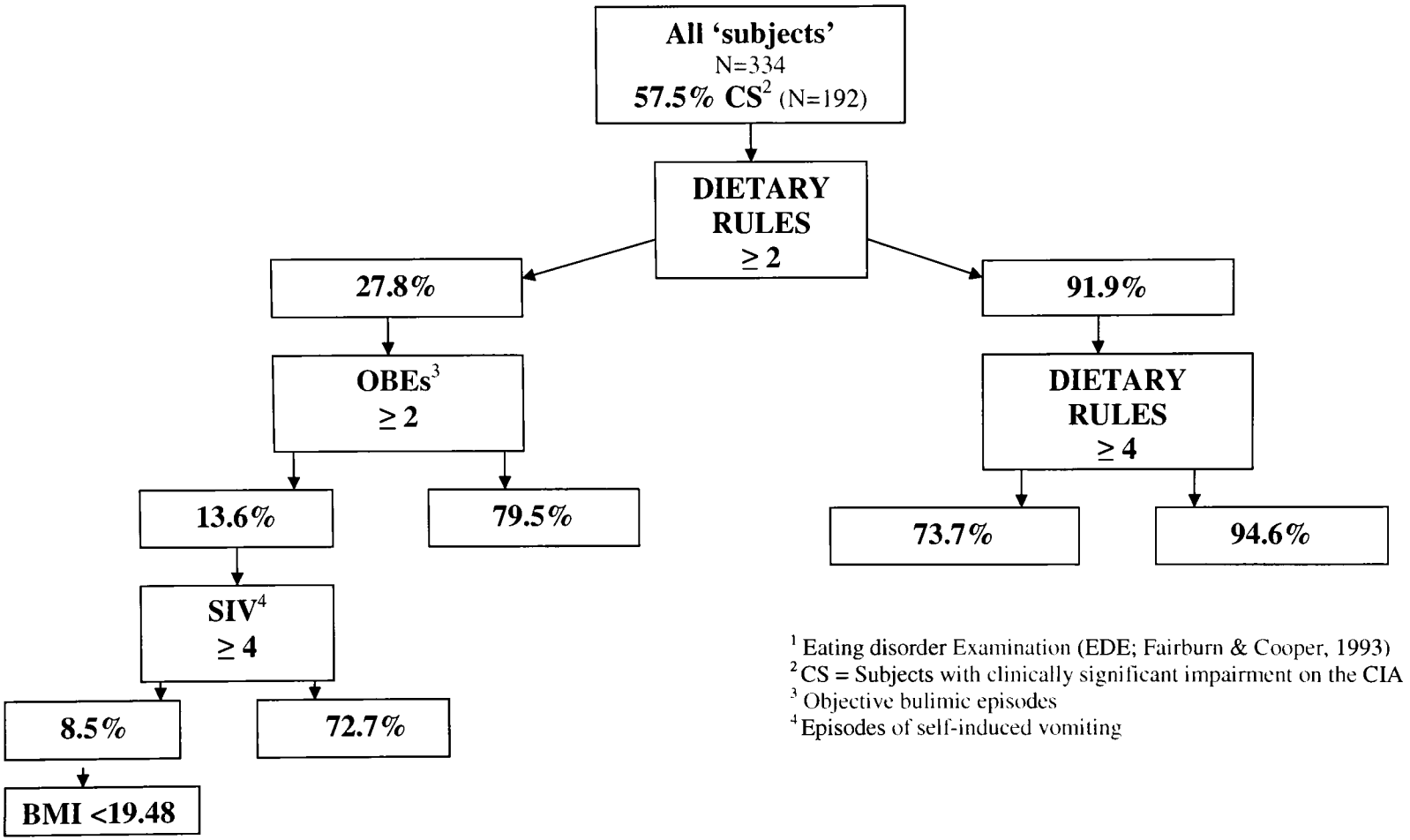
EDE items included in the discriminant function	**Variable loading**	**Standardised coefficient**
Dietary rules	0.774	0.585
Importance of strict control	0.631	0.253
Importance of weight and/or shape	0.601	0.242
Dissatisfaction with shape and/or weight	0.534	0.301
ABE (All bulimic episodes)	0.299	0.273

Signal detection analysis

In this analysis, 'subjects' with missing data were excluded on a node-by-variable basis. Between 326 (97.6%; Importance of strict control) and 334 (100%; Purging) of the 334 'subjects' were included in the analysis.

Four of the nine EDE items were selected by the signal detection analysis for being most predictive of the presence of clinically significant impairment. These were dietary rules, objective bulimic episodes, purging episodes and body mass index. The resulting decision tree is shown in Figure 5.3 below. The decision tree combines various predicting variables with 'and/or' rules to best predict the binary outcome, here the presence of clinically significant impairment.

Figure 5.3 *Decision tree created by signal detection analysis combining thresholds on four EDE[1] items so that the presence of a clinically significant level of impairment is best predicted* (NB: the percentages present the rates of clin. sig. impaired subjects)



[1] Eating disorder Examination (EDE; Fairburn & Cooper, 1993)
[2] CS = Subjects with clinically significant impairment on the CIA
[3] Objective bulimic episodes
[4] Episodes of self-induced vomiting

It can be seen in Figure 5.3 above that subjects with a rating greater or equal 2 on the dietary rules item had a 91.9% chance of being significantly impaired. This figure rose to 94.6% if they had a rating of greater or equal 4 on the same item. Subjects who scored below 2 on the dietary rules item still had a 79.5% chance of being clinically significantly impaired if they had shown two or more objective bulimic episodes. Subjects who scored below 2 on the dietary rules item and had shown only one or no objective bulimic episodes, but had induced vomiting four or more times had a 72.7% chance of being clinically significantly impaired. All identified cut-points were significant at the $p<0.001$ level, with the exception of a rating of greater or equal 4 on the dietary rules item, which was significant at the $p<0.01$ level. A decision tree like this can be helpful when making clinical decisions. The information that can be drawn from it for the purpose of the present study is that the above mentioned four EDE items, and amongst them the dietary rules item in particular, were selected by signal detection analysis to be most predictive of the presence of a clinically significant level of impairment.

Table 5.19 is a summary table. It shows the EDE items selected for inclusion by the three statistical methods.

Table 5.19 *EDE items selected for inclusion by the three statistical methods in the detailed analysis (in bold: EDE items selected by two or more statistical methods)*

EDE variables	Logistic regression	Linear discriminant Function	Signal detection
Dietary rules	x	X	x
Dissatisfaction with sh/wt	x	X	
Objective binge episodes	x	-	x
All bulimic episodes		X	
Purging episodes	x	-	x
Importance of strict control	x	X	-
Importance of shape/wt		X	
Body Mass Index			x

Five of the nine EDE items were selected by at least two statistical methods. These were dietary rules, dissatisfaction with shape and/or weight, objective bulimic episodes, purging episodes, and importance of strict control over eating. The EDE item dietary rules was the first item selected by each statistical method, which indicates that it was most strongly associated with the presence of clinically significant functional impairment.

5.4.6.3 Determination of the cut-off points on the final EDE items

The last step in the analysis was to identify the levels of severity of the five identified eating disorder features that were most predictive of being clinically significantly impaired.

Logistic regression analysis

Three hundred and twenty-one 'subjects' (96.1%) were included in this final analysis. The five EDE items that had been shown to be most predictive of impairment status were entered. A selection of potential cut-points on each item was identified in advance. For the categorical EDE items these were:

1. a rating of < 2 versus ≥ 2
2. a rating of < 3 versus ≥ 3
3. a rating of < 4 versus ≥ 4, and
4. a rating of < 5 versus ≥ 5.

For the continuous EDE items the following cut-points were pre-selected:

1. < 1 episode versus ≥ 1 episodes/28days
2. < 2 episodes versus ≥ 2 episodes/28days
3. < 3 episodes versus ≥ 3 episodes/28 days
4. < 4 episodes versus ≥ 4 episodes/28 days, and
5. < 8 episodes versus ≥ 8 episodes/28 days.

The EDE items were entered one-by-one using Wald's stepwise forward selection for all potential cut-points. This was done in order to select the cut-points most significantly associated with a clinically significant level of impairment.

Table 5.20 shows the optimal cut-points that logistic regression analysis selected for each EDE item. Their importance in predicting a clinically significant level of functional impairment when considered independently are also shown in the table: the Exp(B) value indicates how many more times a person is likely to be significantly impaired if they present with the eating disorder feature in question at or above the selected threshold, than if they show the feature below the threshold or not at all.

Table 5.20 *Thresholds (on five final EDE items) selected by logistic regression analysis to be significantly associated with impairment and their importance in predicting impairment*

EDE Item and Optimal Cut Point	B	Wald	Df	Sig.	Exp(B)	95% C.I. for Exp(B)	
						Lower	Upper
Dietary rules							
A rating of ≥ 4	2.671	27.524	1	.000	14.461	5.331	39.233
OBE							
≥ 2 episodes/28 days	1.896	17.969	1	.000	6.661	2.772	16.009
Import. of strict ctrl							
A rating of ≥ 3	1.762	13.906	1	.000	5.822	2.306	14.694
Purging episodes							
≥ 3 episodes/28 days	1.530	10.708	1	.001	4.620	1.847	11.555
Dissatisf. with sh/wt							
A rating of ≥ 4	1.729	21.372	1	.000	5.633	2.707	11.721

The Omnibus Tests of Model Coefficients showed that this final model had an excellent 'goodness of fit' (χ^2=250.417; df=5; p<0.001). The Cox & Snell R Square and the Nagelkerke R Square values indicated that between 54.2% and 72.6% of the variation in impairment status were explained by the five selected EDE items at their optimal cut-points. The model correctly classified 89.4% of the 321 subjects. The sensitivity and specificity of the model were 90.6% and 87.9% respectively.

Applying the findings to the study dataset

Table 5.21 shows the number of 'subjects' who scored above the identified cut points on the five EDE items. It can be seen that 24 of the 321 'subjects' (7.5%) had a "positive" score (here defined as 'above the cut point') on all five EDE items, 39 (12.2%) scored

positively on four EDE items, 61 (19.0%) on three EDE items, 49 (15.3%) on two EDE items, 51 (15.9%) had a positive score on one EDE item, and 97 (30.2%) did not score above cut point on any of the 5 EDE items.

The table also indicates that 87.8% of all subjects in the CS group had high scores (i.e., above cut point) on two or more of the five EDE items, whereas this was only true for 10% of the subjects in the NS group. The remaining 90% of the NS group either scored below the cut point on all (65.7%) or all bar one (24.3%) of the five EDE items.

Table 5.21 *Number of subjects who scored above the optimal cut points on either none, one, two, three, four, or all of the identified EDE items*

Number of "positive" EDE items	**NS group[1] (N=140)**	**CS group[2] (N=181)**	**Total (N=321)**
0	92	5	97
% within this group	94.8%	5.2%	100.0%
1	34	17	51
% within this group	66.7%	33.3%	100.0%
2	9	40	49
% within this group	18.4%	81.6%	100.0%
3	4	57	61
% within this group	6.6%	93.4%	100.0%
4	1	38	39
% within this group	2.6%	97.4%	100.0%
5	0	24	24
% within this group	0%	100%	100%
Total	**140**	**181**	**321**
% within this group	**43.6%**	**56.4%**	**100.0%**

[1] Subjects without clinically significant global impairment on the CIA
[2] Subjects showing clinically significant impairment on the CIA

5.4.7 Summary of findings

- Preliminary analyses indicated that the following nine EDE items showed the strongest association with a clinical level of functional impairment: Dietary rules, Objective bulimic episodes, All bulimic episodes, Purging episodes, Desire to lose weight, Importance of strict control over eating, Dissatisfaction with shape and/or weight, Importance of shape and/or weight and Body mass index.
- Five of these nine EDE items were selected by at least two out of three independent statistical methods as being most strongly predictive of the presence of clinically significant impairment. These were:

  Dietary rules
  Objective bulimic episodes
  Purging episodes
  Dissatisfaction with shape and/or weight
  Importance of strict control over eating

- A logistic regression analysis identified optimal 'cut points' for these five EDE items, at or above which the likelihood of the presence of a clinical level of impairment was statistically significantly increased. These were (for a time period of 28 days):

  The pursuit of definite rules regarding one's eating on at least 16 days
  The presence of at least two objective bulimic episodes
  The presence of at least three purging episodes
  Moderate to marked dissatisfaction with one's shape and/or weight (i.e., a rating of 4 or more on this EDE item)
  Moderate to supreme importance of maintaining strict control over one's eating (i.e., a rating of 3 or more on this EDE item).

- The model correctly classified 89.4% of the 321 subjects as being clinically significantly impaired. The sensitivity and specificity of the model were 90.6% and 87.9% respectively.
- Applying these findings to the study dataset revealed that the presence of any two or more of the identified five EDE items with a rating at or above their optimal cut point best discriminated the subjects with clinically significant impairment from those without.

5.5 Discussion

Studies of the Reliability of the CIA

Two studies were conducted to examine the reliability of the CIA. The first tested inter-rater reliability and the second test-retest reliability. Both designs involved the administration of the EDE and the CIA (by a trained research assistant) to a group of participants with varying degrees of eating disorder psychopathology. The groups differed for each study. The design of the inter-rater reliability study involved the candidate listening to and rating the tape recordings of these interviews. The test-retest reliability study involved the candidate re-administering the interview after three-to-seven days and making her own ratings. In both studies the ratings for each CIA domain and global impairment were compared.

The inter-rater reliability for the six CIA domains was high with 83.9% to 96.8% exact agreement in the ratings of the two independent raters. The kappa values indicated that agreement was substantial-to-outstanding for five of the six CIA domains, and moderate for the domain Mood and Self-Perception. The global impairment rating showed good inter-rater reliability with 87.1% exact agreement between the two raters.

The test-retest reliability for four of the six CIA domains was excellent with 90.0% to 100% exact agreement in the ratings of the two raters. The two domains Mood and Self-Perception and Interpersonal Impairment showed moderate-to-good test-retest reliability with the percentage exact agreement between ratings at the two time points being 63.6% and 81.8% respectively. The global impairment rating also showed good inter-rater reliability with 72.7% exact agreement between the two interviews.

A strength of both studies was the nature of the sample assessed. The participants in both studies had a broad range of eating disorder psychopathology. Therefore the full scoring range of the CIA was tested (from mild to severe impairment). However, the sample was also problematic in some respects. It was ethnically homogeneous and small in size, especially in the test-retest reliability study. Initially it was planned to assess a larger sample for the test-retest reliability study, but several obstacles prevented this from happening: first, the candidate could not assess any patients she had treated herself; second, many patients no longer showed eating disorder features at their end-of-treatment and/or follow-up assessments, and were therefore not administered the CIA; and third, some eligible patients declined to take part. Further reliability studies using larger and more heterogeneous samples are therefore needed.

Another limitation concerns the design chosen for the two studies. The design for the inter-rater reliability study was the conventional one but it was not ideal since it only tested one aspect of inter-rater reliability. In an investigator-based interview like the CIA interviewers are merely given guidance with regard to exactly what questions to ask and how to phrase them (i.e., how to elicit the information being sought). The ratings are then made on the basis of participants' responses and on further questioning by the interviewer. Therefore with an instrument like this, there are two sources of variability between two

raters assessing the same person: the nature of the questioning and the way the information elicited is rated. The current study tested only the rating aspect (as questioning was only done by the initial interviewer). It did not test the way information was elicited. This design was nevertheless chosen since the other option was not practicable; namely, conducting two interviews in quick succession using two different raters. The latter design would have imposed too much of a burden for participants and, in any case, the participant would remember the first interview and this would be likely to affect his or her responses to the second one.

The design of the test retest reliability study was also not ideal. The use of two different raters meant that the design did not only test the stability of CIA ratings over time but it also tested it across two different raters. To get around this confound, both interviews could have been conducted by the same interviewer. The shortcoming of this design is that it is highly likely that the interviewer would remember the first interview and this would affect the conduct of the second one. The current design is therefore to be preferred but it is a demanding one since it involves two sources of error – the passage of time and a change of interviewer.

Taking these limitations into consideration, it can be concluded that, overall, the CIA showed good inter-rater and test-retest reliability. The test-retest findings are, however, based on a very small sample. It therefore needs to be repeated using a larger one. It is of note that in both studies the Mood and Self-Perception domain was the one that had the least number of exact agreements between the two raters. It seems that it is relatively hard for assessors to judge the presence and severity of secondary impairment in mood and the way that participants feel about themselves. This is probably because these features are inherently difficult to elicit and evaluate, and it can be unclear whether any abnormalities

found are truly secondary to eating disorder features. Further refinement of this section of the CIA is therefore warranted.

Studies of the Validity of the CIA

Both the construct and discriminant validity of the CIA were studied. Testing the construct validity of the CIA involved examining whether the CIA was a valid measure of both the clinical significance and the severity of psychosocial and physical impairment secondary to eating disorder psychopathology. Ratings on the CIA (administered by a trained research assistant) were compared with those of an expert clinician, the ratings being of the same participants and covering the same period of time. Three specific comparisons were made.

The discriminant validity of the CIA was tested in two ways. The first test was to examine whether there were more participants with an eating disorder of clinical severity in the group of participants that had been classed as clinically significantly impaired using the CIA than in a group that was not judged to be clinically significantly impaired using the CIA. The second test involved comparing the total CIA severity scores of patients with an eating disorder of clinical severity with those of participants who were judged to no longer have a clinical eating disorder.

With regard to the three tests of construct validity, it was found that all three comparisons suggested that the CIA was a valid measure of the clinical significance and the severity of secondary functional impairment. The expert clinicians' impairment ratings were statistically significantly higher for participants who had been classed as being significantly impaired by the CIA than for those who had not been so classed. They also correlated strongly with the total severity scores of the CIA. The expert clinicians' impairment ratings also matched the three severity categories created from ratings on the CIA with participants

in the 'severe impairment' group receiving the highest impairment ratings, followed by the participants in the 'moderate' group, and then those in the 'minimal' impairment group.

With regard to the two tests of discriminant validity, the CIA discriminated well between participants with and without an eating disorder. About eighty percent of the participants who were judged to be clinically significantly impaired on the CIA had an eating disorder of clinical severity. This was true of only 7.5% of the participants who did not show significant impairment on the CIA. Similarly, the total severity scores on the CIA were statistically significantly higher in participants with an eating disorder of clinical severity than in those without one.

A potential limitation of the construct validity study was the choice of validator. Ideally the performance of the newly developed interview would have been compared with that of an already well-established measure of functional impairment secondary to eating disorder psychopathology. However, this was not possible as such a measure does not exist. It was therefore decided to test the construct validity of the CIA by comparing its performance against the clinical way of assessing the impact that an eating disorder is having on a person's life: the judgement of an expert clinician. An expert clinician can provide a direct assessment of the construct that the CIA is intended to measure, the severity and clinical significance of psychosocial and physical impairment secondary to eating disorder psychopathology. A related limitation is the fact that only two senior clinicians took part. It would have been preferable to use additional expert clinicians but this was not feasible.

The choice of validator for the discriminant validity study was also not ideal but it was the only one possible given the lack of an existing measure of impairment. The reason for choosing 'eating disorder case status' as the validator was that it can be assumed that the level and severity of secondary impairment is higher in people with an eating disorder than

in those without (given the definition of a psychiatric disorder - see Chapter One, Section 1.3.3).

A major strength of these studies was the nature of the sample. The participants had a broad range of eating disorder psychopathology and secondary impairment, which meant that the full scoring range of the CIA was tested. Another strength was the use of the EDE, a well-established measure for characterising eating disorder features. A weakness of the sample was the fact that is was largely ethnically homogeneous. It cannot be assumed that the same findings would emerge from studies of other ethnic groups. A further weakness of the sample was that it contained only few patients with anorexia nervosa. It would be an important step to replicate these findings with a different eating disorder sample. Lastly, the use of sub-samples for the individual studies was not ideal.

In summary, taking the above mentioned limitations in consideration, the findings of this study suggest that the CIA is a valid measure of the severity and clinical significance of psychosocial and physical impairment secondary to eating disorder psychopathology. Ratings on the instrument matched the judgement of two independent expert eating disorder clinicians. The CIA also possessed good discriminant validity in that it successfully discriminated between participants with a clinical eating disorder and those without.

Identifying those eating disorder features that are associated with clinically significant secondary psychosocial impairment

The study had two aims. The first was to identify those features of an eating disorder that are associated with a clinical level of functional impairment, and the second was to determine the level of severity of these features that is best predictive of being clinically significantly impaired. The design involved administering the EDE and the CIA to a group of subjects exhibiting the complete spectrum of severity of eating disorder disturbance.

In a preliminary analysis the eating disorder features (i.e., EDE items) of participants who were classed on the CIA as clinically significantly impaired were compared with those of participants who did not show clinically significant impairment. It was found that the great majority of EDE items (25/32) had a statistically significant association with impairment status and therefore successfully discriminated between the two groups.

An exploratory logistic regression analysis selected nine of the 25 EDE items as being most predictive of being clinically significantly impaired. These nine EDE items (see Table 5.16 above) were entered into three independent multivariate analyses: logistic regression, discriminant function and signal detection analysis. Five of these nine EDE items were selected by at least two out of the three methods as being most strongly predictive of the presence of clinically significant impairment. These were Dietary rules, Objective bulimic episodes, Purging episodes, Dissatisfaction with shape and/or weight, and Importance of strict control over eating. The EDE item 'Dietary rules' was selected first by all three statistical analyses which indicates that it was most predictive of a clinically significant level of functional impairment. A final logistic regression analysis identified optimal 'cut-points' for these five EDE items. The model correctly classified 89.4% of the 321 subjects as being clinically significantly impaired. The sensitivity and specificity of the model were 90.6% and 87.9% respectively.

Applying these cut-points to the current dataset revealed that the presence of any two or more of the identified five EDE items at or above their optimal cut point best discriminated the participants with clinically significant impairment from those without. The great majority (87.8%) of impaired participants had high scores (i.e., at or above the cut point) on two or more of the five EDE items, whereas this was only true for 10% of the participants without impairment.

This study had two important strengths. A major strength was the sample. It consisted of a clinical sample of people with a broad range of eating disorder features in terms of both their nature and their severity. It was also a sample of patients assessed before and after treatment. These are the types of people whose clinical severity (in terms of impairment) and case status frequently have to be determined by clinicians. Another strength was its use of investigator-based measures of eating disorder features and secondary impairment.

A limitation of the study was that its findings were based on the assumption that the CIA (a new instrument) is a good measure of clinically significant secondary functional impairment. Having said this, the CIA proved to possess good construct and discriminant validity (see section 5.3.3 above). Furthermore, no other measure of clinically significant impairment due to eating disorder features exists. As pointed out above, it would be interesting and important to replicate the construct validity of the CIA using different expert eating disorder clinicians. The next step would then be to replicate the findings of the current study using a different sample.

Another limitation of the study was that all available EDE and CIA data of participants were used, rather than only one set of data per participant. The reason for doing this was to increase the power of the study. However, by doing so, statistical independence was violated. Since the preliminary analysis revealed that the selected EDE items showed consistent association with impairment over time (i.e., at different assessment times), it can however be assumed that the EDE and CIA responses of participants who had been interviewed multiple times did not differ significantly from the responses of those who had only been interviewed once.

A further limitation of the study is that its findings were derived from an ethnically homogeneous sample. It cannot be assumed that the same findings would emerge from more heterogeneous samples. Lastly, and as always, the need for replication needs to be stressed.

In summary, five eating disorder features were found to be highly predictive of the presence of a clinically significant level of psychosocial and physical impairment and the threshold level of severity of these features was identified. This means that these five features, when present at or above the threshold levels specified, are highly predictive of the presence of clinically significant impairment in people with eating disorder psychopathology, especially if two or more are rated positively. Given that the presence or absence of secondary clinically significant impairment defines what is, or is not, a "case", these features essentially provide an operational definition of what constitutes an "eating disorder". With reference to Figure 1.1 (see p. 141), these items at their specified threshold levels can be viewed as constituting the boundary of an eating disorder as represented by the outer circular line of the figure.

With regard to clinical practice, on the basis of this study's findings not all five of the eating disorder features need to be present for a person to be classed as a "case": an eating disorder could be said to be present if the person in question has the psychopathology characteristic of an eating disorder and met two or more of the following EDE-based criteria (over the past 28 days):

1. ... Had been attempting to obey strict dietary rules on at least 16 days
2. ... Had experienced at least two objective bulimic episodes
3. ... Had experienced at least three episodes of purging (laxative or diuretic misuse)
4. ... Had been at least moderately dissatisfied with their shape or weight

5. ... Had judged their self-worth to at least a moderate extent on the basis of their ability to maintain strict control over their eating.

Thus a "transdiagnostic" impairment-based definition of caseness has been generated. This contrasts with the largely arbitrary, symptom-based (rather than impairment-based), definitions of anorexia nervosa and bulimia nervosa (see Chapter One, Section 1.3.4). At the same time a definition of eating disorder NOS has (by default) also been created; namely those cases of an eating disorder (as defined above) that do not meet diagnostic criteria for anorexia nervosa or bulimia nervosa.

Were this study's findings found to be robust on the basis of further studies using these methods or other applicable ones, then clinicians and researchers would have available to them an easily applied definition of what constitutes an eating disorder. This could be used in epidemiological studies (for example, to determine the prevalence of eating disorders including eating disorder NOS) and in clinical research (for example, to define treatment outcome in terms of case status). It would also provide clinicians with a simple clinically-meaningful definition of what is an eating disorder, something that is lacking at present.

CHAPTER SIX
The Development of a Self-report Measure of Impairment – the CIA-Q

6.1 Introduction

As discussed in Chapter One (Section 1.5.2), there is a need for an easily-administered, clinically useful measure of psychosocial impairment secondary to eating disorder psychopathology. To date there is no satisfactory measure of this type. The measures that exist have a number of shortcomings. These include not ensuring that the impairment is truly secondary to eating disorder features (rather than say clinical depression) or confounding the measurement of impairment with the measurement of symptoms. The one instrument that does measure true secondary impairment (Engel et al., 2006) neglects to assess impairment due to concerns about shape, a central element of the core psychopathology of eating disorders and possibly the most disabling.

It would be of great value for clinical work and research to have a measure of the severity of secondary psychosocial impairment to complement the existing measures of eating disorder features. Such a measure could have several purposes. First and foremost, it would standardise the way clinicians assess the impact an eating disorder is having on an individual's life. This assessment is necessary for clinicians to determine whether any eating disorder features present result in clinically significant distress or impairment of functioning, a clinical judgment required to make the diagnosis of an eating disorder and to gauge the patient's need for treatment. Second, it could be used to measure outcome after treatment. Simply measuring eating disorder symptoms does not give a complete picture of the state of an individual. It is important to capture the extent to which any symptoms present are interfering with the individual's life. Third, the measure could be used to monitor change over time.

This chapter describes the development and validation of a new measure of psychosocial impairment secondary to eating disorder features, the "Clinical Impairment Assessment – Questionnaire" or CIA-Q. In section 6.2 a detailed account of the development of the measure is given. Section 6.3 describes four studies examining the reliability and the validity of the instrument. In section 6.4 the findings are summarised and discussed.

6.2 The Development of the CIA-Q

In this section of the chapter the major steps in the development of the Clinical Impairment Assessment - Questionnaire (CIA-Q) are outlined. A detailed description of the current version of the CIA-Q (CIA-Q version 1.0) is also given. It can be seen in Appendix 6.1. The first step in the development of the CIA-Q was to decide which features the instrument required. Given its purpose, it was decided that it should have the following characteristics:

- Brief self-report questionnaire – to ensure that the CIA-Q could be easily administered[13];
- Should focus on the participant's state over the past 28 days – to ensure that the CIA-Q assessed the current severity of impairment;
- Should generate a total score – to provide a simple index of severity;
- Should focus exclusively on domains of life that are affected by eating disorder features – to ensure the relevance of its content;

[13] A self-report measure seemed a satisfactory method for obtaining data about secondary psychosocial impairment since, unlike the psychopathology of eating disorders, the phenomena are not conceptually complex nor do professional and lay concepts differ.

- Participants should consider all aspects of their eating disorder psychopathology when rating how their life has been affected – to ensure that the full extent of secondary impairment is rated;
- Should be administered immediately after the assessment of current eating disorder psychopathology (using a self-report instrument such as the Eating Disorder Examination-Questionnaire [EDE-Q; Fairburn & Beglin, 1994]) – to ensure that any eating disorder features present are brought to the forefront of the participant's mind before he or she answers questions about how these features have been interfering with his or her life;
- Should be designed to ensure that the impairment rated is secondary to eating disorder features – to avoid rating impairment due to other causes (e.g., comorbid clinical depression).

The next step in the development of the CIA-Q involved identifying the domains of life most prone to be affected by eating disorder features. This was done in collaboration with three eating disorder experts (Dr Zafra Cooper, Professor Christopher Fairburn, and Professor Robert Palmer), each of whom had been working in this field for at least two decades and had extensive experience assessing and treating patients with eating disorders. The Clinical Impairment Assessment – interview version (CIA; see Chapter Five) had already been developed and implemented with eating disorder patients at the time of the development of the CIA-Q. Therefore the domains of functioning primarily affected by eating disorder psychopathology were already well-established (see Chapter Five, Section 5.2.1). It was decided to adopt four of the six CIA domains: Mood and Self-perception, Cognitive effects, Interpersonal impairment, and Work impairment. It was decided not to include the domain 'Physical Health' since it was the collective opinion of the candidate and

the three experts that it would be impossible for patients to judge the full extent to which their eating and weight-control behaviour were affecting their physical functioning. Differences between lay and professional concepts with regard to physical functioning might have further complicated the self-report assessment of this domain. The domain 'Life not otherwise specified (NOS)' (see Chapter Five, Section 5.2.5) was also not included since it would have been difficult to standardise it.

The next step involved thinking of the content of each domain of functioning (i.e., the way that each domain could be affected by eating disorder psychopathology). The candidate and the three experts each listed examples of impairment due to eating disorder psychopathology for each identified domain. They were guided by their own clinical experience and by the answers given by patients who had been interviewed using the CIA. In addition, generic Health Related Quality of Life (HRQOL) instruments were reviewed and searched for areas of potential impairment that had not already been included (Boini et al., 2004; Kolotkin et al., 2001; Spilker & Revicki, 1996). The result was the generation of a list of domain-specific ways in which eating disorder psychopathology might affect psychosocial functioning (see Table 6.1). This list was carefully evaluated by the candidate and the experts, and examples of impairment that seemed redundant, highly unusual, or were overlapping with other examples were removed. The final list consisted of 22 forms of impairment (see Table 6.1; examples marked with *), each of which was chosen to be an item for the CIA-Q. Seven were examples of impairment regarding mood and self-perception, four were examples of cognitive effects, seven were examples of interpersonal impairment, and four were forms of work impairment.

Table 6.1 *Examples of psychosocial impairment secondary to eating disorder features*

Domain of Functioning	**Examples of impairment due to eating habits and/or concerns about eating, shape and weight**
Mood and Self-Perception	feeling a failure *
	· feeling hopeless
	· feeling sad
	being tearful
	being irritable *
	· feeling guilty *
	lacking self-confidence
	· feeling angry
	worrying *
	feeling revolted by oneself
	feeling self-critical *
	feeling disgusted with oneself
	feeling overwhelmed
	· being ashamed *
	being upset *
Cognitive Functioning	· having difficulties concentrating *
	mind drifting off frequently when reading
	being forgetful *
	having difficulty making everyday decisions *
	being preoccupied with thoughts about food, shape or weight
	· being absent-minded *
	· not being able to focus on conversations
Interpersonal Functioning	impaired socialising *
	· feeling uncomfortable around people
	difficulty eating with friends or relatives *
	impaired sex life *
	not wanting to be touched by partner
	not wanting to be seen naked
	difficulty eating out with others *
	arguments about eating behaviour or weight with family/partner/friends *
	difficulty feeding children
	· not doing things that used to be enjoyed *
	· interference with relationships with others *
Work Impairment	· impaired performance at work *
	having to take time off work *
	finding it more difficult to work *
	spending less time on work *

* These examples were adopted into the final version of the CIA-Q

The next step was to generate questions that addressed each of these forms of impairment. Thus 22 questions were created. Each question started with: "Over the past month, to what extent have your eating habits, exercising or feelings about your eating, shape or weight affected"... (the domain of impairment in question). A Likert rating system was used with the response options being 'Not at all', 'A little', 'Quite a bit', and 'A lot'. These responses were scored 0 to 3 respectively, with a higher rating indicating a higher level of impairment. A Likert rating system was chosen since it works well with the widely used self-report measure of eating disorder psychopathology, the Eating Disorder Examination - Questionnaire (EDE-Q; Fairburn & Beglin, 1994), the measure that was identified as the natural partner of the CIA-Q[14]. For a review of the EDE-Q, see below (end of this section).

The time frame of the EDE-Q (the past 28 days) was adopted for the CIA-Q. The 22 items were presented in random order (i.e., they were not presented in a domain-specific way) to encourage more accurate reporting.

To obtain an overall index of severity the ratings on the 22 items were averaged, the resulting figure (ranging from 0 to 3) being termed the 'Global Impairment Index'. Thus a higher Global Impairment Index was designed to indicate a greater level of secondary psychosocial impairment.[15]

The CIA-Q was then piloted on 33 non-eating disordered subjects and a small sample of patients with an eating disorder (n=10) who were currently receiving treatment. On the basis of their feedback, a number of minor modifications were made to the wording of the questions. The final version of the CIA-Q is shown in Appendix 6.1.

---

[14] As mentioned above, it was decided that the CIA-Q should always be preceded by a self-report measure of eating disorder psychopathology, in order for participants to have their eating disorder features 'fresh in their mind' when answering questions about the impact these features were having on their everyday life.

[15] It was decided to not calculate domain-specific scores since the purpose of the instrument was to measure the overall severity of secondary psychosocial impairment, and not the impairment in specific domains of life.

The partner instrument that was chosen for the CIA-Q was the Eating Disorder Examination – Questionnaire (EDE-Q; Fairburn & Beglin, 1994). This instrument was adapted from the interview version of the EDE (Fairburn & Cooper, 1993) (see Chapter Three, Section 3.2.2), which has considerable reliability and validity data to support its use. Although the two versions of the EDE contain identical items, a number of investigators have found that the agreement between the questionnaire and the interview version is inconsistent (Black & Wilson, 1996; Fairburn & Beglin, 1994; Wilfley, Schwartz, Spurrell & Fairburn, 1997), especially for more complicated constructs like binge eating. Preliminary data indicate that the EDE-Q has acceptable levels of reliability (Luce & Crowther, 1999) but further investigations are needed to establish the psychometric properties of the instrument. The 33-item EDE-Q focuses on eating disorder symptoms and attitudes over the past 28 days and contains four subscales: Restraint, Eating Concern, Shape Concern and Weight Concern. Most items are rated on a 7-point scale (0-6), and raw 28-day frequencies are used for the behavioural items. The Global EDE-Q score is the average of the four subscale scores and is an index of the overall severity of eating disorder psychopathology. The current version of the EDE-Q (version 5.2) is shown in Appendix 6.2.

6.3 The Reliability and Validity of the CIA-Q

Four studies were conducted to examine the reliability and the validity of the CIA-Q. The goal of the first was to test the internal consistency and the factorial structure of the instrument. The goal of the second was to establish its test-retest reliability. The goal of the third was to test its construct and discriminant validity. The goal of the fourth was to establish whether the instrument was sensitive to change (i.e., was responsive). The results

of each of these studies are summarised at the end of each study. The studies as a whole are discussed in the overall discussion in Section 6.4.

The participants for the four studies were subgroups of the 170 patients described in Chapter Three. As outlined in Chapter Three, these patients were consecutive referrals to one or other of two outpatient eating disorder clinics in the UK (Oxford and Leicester) for the treatment of an eating disorder. The diagnostic distribution of the total sample at the time of referral was as follows: 8 patients (4.7%) fulfilled diagnostic criteria for anorexia nervosa, 60 (35.3%) were diagnosed with bulimia nervosa, and 102 (60.0%) received the diagnosis eating disorder NOS. As part of the research assessments at the beginning and end of their treatment, and at 20-, 40-, and 60-week post-treatment follow-up, and at two and three years after treatment, they were asked to complete the EDE-Q immediately followed by the CIA-Q. They were also seen by one of the three local expert clinicians (Dr Zafra Cooper or Professor Christopher Fairburn in Oxford; Professor Robert Palmer in Leicester) at each of these times, who determined whether the participant had a clinical eating disorder (see Chapter Three, Section 3.2.2). In Oxford, the two expert clinicians additionally assessed to what extent the participant's eating habits and their concerns about eating, shape or weight had influenced their psychosocial functioning over the previous 28 days (i.e., the same time frame as the completed CIA-Q) (see below).

There were three main reasons for the use of subgroups of patients rather than the whole sample: first, some of the studies were only conducted at one of the two eating disorder clinics, namely the one in Oxford. This reduced the number of potential participants to 105. Second, the CIA-Q was only introduced to the research protocol after 120 patients had already started or completed treatment. Third, the CIA-Q data of patients who were diagnosed by Professor Fairburn as suffering from a severe comorbid clinical

depression at the time of assessment were excluded since some of the impairment reported by such patients might have been secondary to the clinical depression rather than to the eating disorder features present.

The data used in these studies are taken not only from the first assessment (i.e., before treatment), but also from the end of treatment and from the follow-up times. This was done since it seemed crucial to validate the CIA-Q using a sample of individuals exhibiting the full range of eating disorder psychopathology, and therefore the full range of secondary impairment.

6.3.1 Study One: Internal consistency and factor structure of the CIA-Q

Design

This study had two aims. The first was to test one aspect of reliability of the newly developed instrument, its internal consistency (i.e., the degree to which the individual items measure the same underlying construct). The second was to examine the factorial structure of the instrument. This was done in order to test whether the CIA-Q was unidimensional (i.e., measuring only one underlying construct).

Participants

The sample consisted of 127 of the 170 patients described in Chapter Three. These were the number of patients who had completed the CIA-Q at least once. The dataset was constructed by using all 'first-time' CIA-Q data[16] (i.e., the CIA-Q data from the assessment point at which the participant first completed a CIA-Q – this was not necessarily at the start of treatment, as explained above). The data distribution for this study was as follows: data from the beginning of treatment were used for 50 participants, from the end of treatment for 32, from 20-week follow-up for 9, from 40-week follow-up for 20, from 60-week follow-up

[16] This was done in order to ensure that no participant had completed the CIA-Q more than once. In contrast to the CIA interview (see Chapter Five), taking the 'first-time' CIA-Q data of participants ensured a wide range of eating disorder features and associated impairment, since they were from all assessment times.

for 11, and from 2-year follow-up for 5 participants. Seventy-two of these 127 participants (56.7%) were judged by Professor Fairburn or Dr Cooper to have an eating disorder of clinical severity at the time that they completed the CIA-Q. The other 55 (43.3%) participants were judged no longer to have a clinical eating disorder.

Data analysis

In order to test the internal consistency of the CIA-Q, one of its most commonly used indicators, the Cronbach's alpha coefficient, was calculated. Ideally, the Cronbach alpha coefficient of a scale should be above .7 (Pallant, 2005). In addition, item-total correlations were performed in order to examine the extent to which each item correlated with the total CIA-Q score, the Global Impairment Index. Low values (less than .3) indicate that the item is measuring something different from the scale as a whole. To test whether the CIA-Q was unidimensional (i.e., it only measured one construct), a principal component analysis was performed. Data analyses were performed using SPSS version 14.0.

Principal component analysis

Principal component analysis is a factor analytic technique that can be used to detect structure in the relationship between variables. It takes a large set of variables and looks for a way that the data may be reduced or summarised using a smaller set of factors or components. It does this by looking for groups among the inter-correlations of a set of variables in a way that accounts for most of the variability in the pattern of correlations. There are three main steps in conducting a principal component analysis. The first is to assess the suitability of the data for principal component analysis. This involves examining the sample size and the strength of the relationship among the variables (or items). With regard to the sample size, Tabachnik and Fidell (2001) state that a small sample size (e.g. 150 cases) should be sufficient if solutions have several high loading marker variables

(above .80). With regard to the strength of relationship among the variables they recommend that some correlation coefficients between variables should be greater than .3. Two statistical measures (generated by SPSS) also help assess the factorability of the data: Bartlett's test of sphericity (Bartlett, 1954) and the Kaiser-Meyer-Olkin measure of sampling adequacy (Kaiser, 1974). The first should be significant ($p<.05$) and the latter above .6 for the factor analysis to be considered appropriate (Tabachnik and Fidell, 2001). The second step in conducting a principal component analysis is factor extraction and involves determining the smallest number of factors that can be used to best represent the interrelations among the set of variables. For the purpose of this study, Kaiser's criterion and Catell's scree test (Catell, 1966) were used to aid with the decision concerning the number of factors to retain. Using Kaiser's criterion, only factors with an eigenvalue[17] of 1.0 and above are retained for further investigation. Cattell's scree test involves plotting each of the eigenvalues of all factors and only retaining those factors that are above the 'elbow' (i.e., a break in the plot), as these factors contribute the most to the explanation of the variance in the data set (Cattell, 1966). The third step is factor rotation and the interpretation of the resulting structure. Factor rotation assists in the interpretation of the factors by presenting the pattern of loadings in a way that is easier to interpret. Both orthogonal (Varimax) and oblique (Direct Oblimin) factor rotations were performed. Oblique rotation allows for the factors to be correlated (Tabachnik and Fiddle, 2001).

Results

Internal consistency

The internal consistency of the CIA-Q proved to be high, with a Cronbach alpha coefficient of .96. As can be seen in Table 6.2, all 22 items correlated positively with the Global Impairment Index of the CIA-Q. Item-total correlations ranged from .48 to .86.

[17] The eigenvalue of a factor represents the amount of the total variance explained by that factor.

Table 6.2 Correlations of each individual item with the total CIA-Q score[1]

Item	Corrected Item-Total Correlation
Over the past month, to what extent have your eating habits, exercising, or feelings about your eating, shape or weight ...	
... made you feel a failure	.821
... made you irritable	.742
... made you feel guilty	.811
... made you worry	.798
... made you feel critical of yourself	.806
... made you feel ashamed of yourself	.836
... made you upset	.857
... made it difficult to concentrate	.821
... made you forgetful	.676
... affected your ability to make everyday decisions	.751
... made you absent-minded	.717
... stopped you going out with others	.693
... interfered with meals with family and friends	.723
... affected your sex life (if applicable)	.481
... made it difficult to eat out with others	.720
... led to arguments with others	.591
... interfered with you doing things you used to enjoy	.760
... interfered with your relationships with others	.806
... affected your performance at work (if applicable)	.666
... made you take time off work (if applicable)	.504
... made it more difficult to work (if applicable)	.636
... made you spend less time on work (if applicable)	.567

[1] Global Impairment Index

Principal component analysis

Prior to performing the principal component analysis, the suitability of the data for a factor analysis was assessed. Inspection of the correlations coefficients for all items (correlation matrix) revealed that the majority of coefficients were above .3. The Kaiser-Meyer-Oklin value was .92, exceeding the recommended value of .6 (Kaiser, 1974) and the Bartlett's Test of Sphericity (Bartlett, 1954) reached statistical significance, supporting the factorability of the correlation matrix.

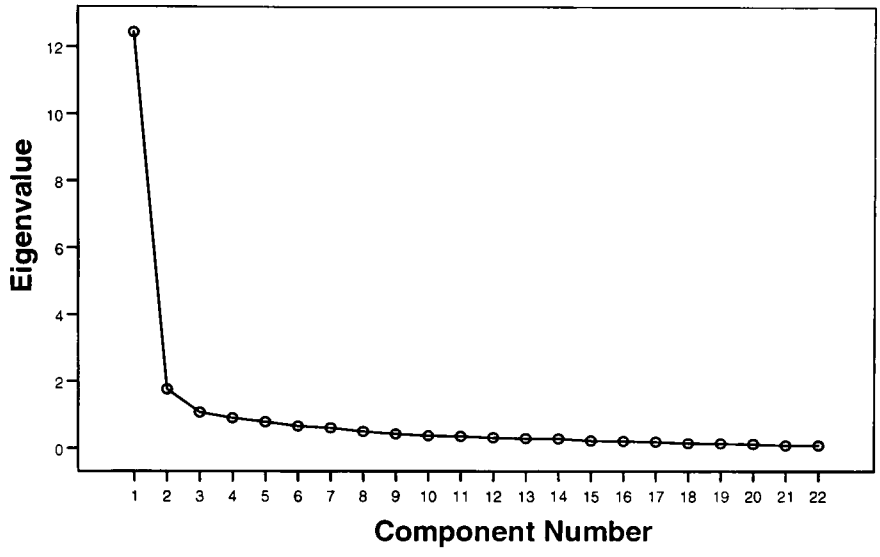
The principal component analysis revealed the presence of three components with eigenvalues exceeding 1, explaining 56.5 per cent, 8.0 per cent and 4.9 per cent of the variance respectively. The component matrix revealed that all items loaded highly on the first component (see Table 6.3). Four items simultaneously loaded on component two, and these loadings were far weaker than on component one.

Table 6.3 *Component matrix showing unrotated loadings of each item on the three extracted components*

Item	**Component**		
Over the past month, to what extent have your eating habits, exercising, or feelings about your eating, shape or weight ...	**1**	**2**	**3**
... made you upset	.876		
... made you feel ashamed of yourself	.858		.357
... made you feel a failure	.846		
... made it difficult to concentrate	.843		
... made you feel guilty	.835		
... made you feel critical of yourself	.831		.305
... interfered with your relationship with others	.829		
... made you worry	.825		
... interfered with you doing things you used to enjoy	.785		
... affected your ability to make everyday decisions	.778		
... made you irritable	.773		
... interfered with meals with family and friends	.752		
... made it difficult to eat out with others	.750		
... made you absent-minded	.749		
... stopped you going out with others	.721		
... made you forgetful	.709		.364
... affected your performance at work (if applicable)	.693	.505	
... made it more difficult to work (if applicable)	.664	.552	
... led to arguments with others	.624	-.359	
... made you spend less time on work (if applicable)	.599	.586	
... made you take time off from work (if applicable)	.533	.498	
... affected your sex life (if applicable)	.517	-.475	

Inspection of the screeplot (see Figure 6.1) revealed a clear break after the first component, and another small break after the second component. Using Catell's scree test, it was decided to retain two components for further investigation.

Figure 6.1 *Scree Plot showing Eigenvalue for each component*



To aid in the interpretation of these two components, Oblimin rotation was performed. This was done, since the performance of Varimax rotation had revealed that the two factors were highly correlated which made Oblimin rotation more suitable. The rotated solution revealed the presence of "simple structure" (Thurstone, 1947), with most items loading strongly on the first component (see Table 6.4).

As can be seen in Table 6.4, 18 of the 22 variables loaded strongly on component one, and only four variables loaded on component two. All of the latter items were questions about secondary impairment in work performance. They also all had an 'if applicable' qualifier, meaning that participants should leave these items blank if they were not currently working. Twenty-eight (22.1%) of the 127 participants left the four work items blank. This is likely to have accounted for these items grouping together on a second component. Overall, the findings suggest that the CIA-Q measures only one construct, i.e., is a unidimensional instrument.

Table 6.4 Pattern Matrix: Loadings of each item on two components rotated with Oblimin

Item Over the past month, to what extent have your eating habits, exercising, or feelings about your eating, shape or weight ...	Component 1	Component 2
... made you worry	**.827**	.034
... made you irritable	**.821**	-.036
... interfered with your relationship with others	**.801**	.080
... led to arguments with others	**.792**	-.220
... made it difficult to eat out with others	**.791**	-.027
... made you upset	**.790**	.169
... affected you sex life (if applicable)	**.787**	-.376
... made you feel critical of yourself	**.773**	.123
... interfered with meals with family or friends	**.772**	.004
... made you feel guilty	**.770**	.135
... made you feel ashamed of yourself	**.765**	.177
... made you feel a failure	**.751**	.178
... interfered with you doing things you used to enjoy	**.693**	.171
... stopped you going out with others	**.668**	.111
... affected your ability to make everyday decisions	**.646**	.230
... made you forgetful	**.618**	.167
... made it difficult to concentrate	**.593**	.409
... made you absent-minded	**.529**	.361
... made you spend less time on work (if applicable)	.078	**.800**
... made it more difficult to work (if applicable)	.157	**.781**
... affected your performance at work (if applicable)	.217	**.738**
... made you take time off work (if applicable)	.086	**.687**

Summary of findings

- The CIA-Q showed excellent internal consistency with a Cronbach's Alpha of .96.
- All individual item scores showed a significant positive correlation with the total CIA-Q score, the Global Impairment Index.
- Performance of principal component analysis revealed that 56% of the instrument's variance was explained by the first of three factors exceeding an eigenvalue of 1.0. All items loaded highly on this factor. This was backed up by the scree plot. Oblimin rotation revealed that 18 of the 22 items showed strong loadings on the first factor.

- These findings suggest that the CIA-Q is an internally consistent and unidimensional instrument.

6.3.2 Study Two: Test-retest Reliability of the CIA-Q

Design

The aim of this study was to evaluate the test-retest reliability of the CIA-Q, that is the stability of ratings on the CIA-Q. The design involved administering the EDE-Q immediately followed by the CIA-Q on two occasions within three days to a group of subjects with varying degrees of eating disorder psychopathology (i.e., exhibiting the complete spectrum of severity of eating disorder disturbance) and comparing the Global Impairment Indices of the CIA-Q on the two occasions. On each occasion the same 28-day time period was assessed. It was predicted that, if the ratings of the CIA-Q were stable over time, there should be a statistically significant positive association between the Global Impairment Indices at the two time points. In addition, the Global Impairment Indices should not be statistically or clinically significantly different.

Participants

This study was conducted in Oxford. The sample consisted of 43 of the 105 Oxford patients, described in detail in Chapter Three. The data for this study were acquired between June 2004 and July 2005 and were from patients who were attending either their end-of-treatment or one of their post-treatment assessments[18] (at which they had filled in an EDE-Q and a CIA-Q). Each patient was invited to participate in this study if they had not suffered from a comorbid clinical depression in the month before attending their assessment[19].

[18] It was decided to not include participants who were about to start treatment for their eating disorder, as it seemed too demanding (in terms of their engagement in treatment) to ask them to fill in the EDE-Q and the CIA-Q twice.

[19] This was assessed by Professor Fairburn in a clinical assessment that took place at each assessment time point.

Patients were not asked more than once to participate in this study (i.e., at more than one assessment time).

The candidate explained to the patient that the study was a research project that was entirely separate from the routine research assessment that they had just undergone and that taking part in it was voluntary. They were told that it would involve filling in two questionnaires in three days' time and sending them back to the candidate in a stamped addressed envelope. If they agreed, they were given an EDE-Q and CIA-Q in an envelope. Written instructions requested them to fill in the two questionnaires three days later, considering the 28 days prior to their original assessment (i.e., ignoring the intervening three days). This was done so that participants had the same time period in mind as at the original time when rating their impairment on the CIA-Q. Fourty-three patients returned their questionnaires.

These 43 participants took part in this study at the following times: 22 at the end of treatment, eight at 20-week follow-up, two at 40-week follow-up, four at 60-week-follow-up, six at 2-year follow-up, and one at their 3-year follow-up assessment. Nine of the 43 participants (20.9%) were judged to have a clinical eating disorder at the time they took part in this study; the remaining 34 participants (79.1%) were judged no longer to have one.

Measures

The EDE-Q and the CIA-Q have been described in detail above (both in Section 6.2 of this chapter). The re-test version of the CIA-Q was identical to the standard CIA-Q in terms of its items except that the order of the items was changed in order to minimise memory effects. The re-test version of the CIA-Q is shown in Appendix 6.3.

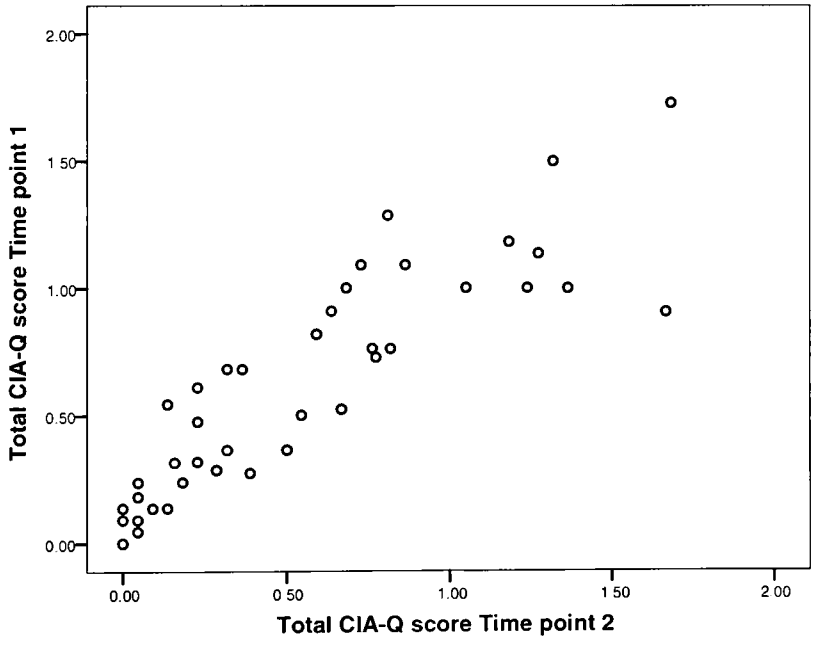
Data analysis

To assess the agreement between the Global Impairment Indices at the two time points, an intra-class correlation coefficient was calculated. In addition, a paired samples t-test was performed to examine the difference between the Global Impairment Indices at the two time points.

Results

The mean Global Impairment Indices at the two time points were 0.59 (SD=0.44) and 0.53 (SD=0.48) respectively. The test-retest intraclass correlation coefficient for the Global Impairment Indices at the two time points was .88 ($p<.001$).

Figure 6.2 *Test-retest reliability: Scatter plot showing the relationship between CIA-Q GII[1] at time point 1 and time point 2*



[1] Global Impairment Index

Figure 6.2 shows a plot of the Global Impairment Index of time point one against time point two. The paired samples t-test revealed that the scores at time one and time two

did not differ significantly (mean difference=0.07, 95% confidence interval=-0.0005 to 0.14, t (42)=2.003, p= 0.052) with the actual difference being only 0.15 of a standard deviation. The difference between the two scores is therefore unlikely to be of clinical relevance. The slightly lower Global Impairment Index at time point two can be explained by regression to mean.

Summary of findings

- A statistically significant positive correlation was found between participants' Global Impairment Indices at the two time points.
- The mean Global Impairment Index at time point two was slightly lower than at time point one. This difference, however, was not statistically significant. The actual difference was very small and is unlikely to be of clinical relevance. It can be explained by regression to mean.
- The findings suggest that the ratings on the CIA-Q are stable over time.

6.3.3 Study Three: Construct and Discriminant Validity of the CIA-Q

Design

The aim of this study was to test the construct and discriminant validity of the CIA-Q. Testing the construct validity of the CIA-Q involved examining whether the CIA-Q was a valid measure of the severity of psychosocial impairment secondary to eating disorder psychopathology. This was done in two ways.

The first test involved comparing the Global Impairment Index with impairment severity ratings of one of the two expert clinicians[20] (Professor Fairburn or Dr Cooper), the ratings being of the same participant and covering the same period of time. The comparison

[20] The clinicians' impairment severity ratings in this study resembled those described in Chapter Five. The only difference was that the clinicians only considered participants' psychosocial impairment when making their ratings (and not also their physical functioning, as in the ratings in Chapter Five).

involved the simultaneous administration of the CIA-Q to, and the clinical assessment of, a group of subjects with varying degrees of eating disorder psychopathology (i.e., exhibiting the complete spectrum of severity of eating disorder features in order to test the full scoring range of the CIA-Q).

The second test was an indirect one. It was based on the assumption that there would be a positive association between the severity of eating disorder psychopathology and the severity of the psychosocial impairment that is secondary to it. The test involved comparing the Global Impairment Index with the severity of associated eating disorder psychopathology as measured by the EDE-Q.

It was predicted that if the CIA-Q was a valid measure of psychosocial impairment due to eating disorder features:

1. There would be a statistically significant correlation between participants' Global Impairment Indices on the CIA-Q and the impairment severity ratings made by the two clinicians,
2. There would be a statistically significant positive correlation between participants' Global Impairment Indices on the CIA-Q and their global EDE-Q score.

The discriminant validity of the CIA-Q was tested by comparing the Global Impairment Indices of participants who were classed by the expert clinicians as having an eating disorder of clinical severity with those of participants who were judged to no longer have a clinical eating disorder. It was predicted that, if the CIA-Q showed discriminant validity, the Global Impairment Indices of participants with a clinical eating disorder would be statistically significantly higher than those of participants without an eating disorder.

Participants

The first test of construct validity was conducted in Oxford. The CIA-Q data of 69 of the 105 Oxford patients described in Chapter Three were used. These were the patients for whom both CIA-Q data and a clinician's rating were available at one or more assessment time points. For the second test of construct validity, the data of 123 of the 170 patients described in Chapter Three were used. These were the patients for whom both CIA-Q data and EDE-Q data were available at one or more assessment time points.

For the test of discriminant validity, the data of 127 patients of the 170 patients were used. These were the patients for whom both CIA-Q data and eating disorder caseness information were available at one or more assessment time points. They were the same participants as in the preliminary study (see section 6.3.1 above).

Measures

The EDE-Q and the CIA-Q have been described in detail above (both in Section 6.2 of this chapter).

Clinicians' rating of secondary psychosocial impairment

The two senior clinicians (Professor Christopher Fairburn and Dr Zafra Cooper) were asked to assess the extent to which the participants' disturbed eating habits and concerns about eating, shape or weight had been having an impact on their life over the previous 28 days. This was done as part of routine assessments of each participant, which took place at the beginning and end of treatment, and at 20-, 40-, and 60-week and 2- and 3-year post treatment. Both senior clinicians were sent an email (see Figure 6.3) by the candidate on the day an assessment was due reminding them to make three ratings: 1) the participant's DSM-IV eating disorder diagnosis over the past month; 2) over the past three months; and 3) the participant's degree of secondary psychosocial impairment over the past month. Figure 6.3

shows the content of this e-mail. The rating of the severity of secondary psychosocial impairment followed the severity scoring system employed by EDE (Fairburn & Cooper, 1993). It consisted of a rating scale of 0 to 6 with higher ratings indicating a more severe level of secondary psychosocial impairment.

The clinicians were asked to make these ratings as soon as possible after their clinical assessment and to send them back to the candidate via email.

Figure 6.3 *Email sent to expert clinicians on the day of the assessment of a patient*

With regard to the forthcoming assessment of the above patient, please would you make the following THREE ratings:

RATING ONE - Eating disorder diagnosis based on DSM-IV criteria over past month
0 – Does not meet DSM criteria for an eating disorder
1 - Anorexia nervosa
2 - Bulimia nervosa
3 - BED
4 - Other EDNOS

RATING TWO - Eating disorder diagnosis based on DSM-IV criteria over last three months

RATING THREE - Impairment over past month
Severity of impairment due to the patient's eating habits, exercising, or feelings about eating, shape or weight OVER THE PAST MONTH.

Consider the following four domains:

- mood, view on self, guilt
- cognitive effects (preoccupation, concentration, forgetfulness)
- interpersonal/social effects (relationships with others, socialising, eating with others, sex)
- work (performance, time off)

0 - No impairment
1
2 - Mild impairment
3 -
4 - Moderate impairment
5 -
6 - Severe impairment

Data analysis

***First test of construct validity**: Relationship between the CIA-Q Global Impairment Index and the clinicians' impairment ratings*

The dataset for this substudy was constructed by using all available and eligible pairs of ratings (CIA-Q Global Impairment Indices and clinicians' impairment ratings) on all of the 69 participants (i.e., from all available assessment times). This meant that each participant 'contributed' one or more pairs of ratings to the dataset. This was done in order to increase the power of the study. The dataset consisted of 142 individual pairs of ratings (i.e., patients' total scores on the CIA-Q and clinicians' ratings of psychosocial impairment). The distribution of these pairs of ratings over the seven assessment time points was as follows: 26 ratings were available from the beginning of treatment, 26 from the end, 23 from 20-week follow-up, 27 from 40-week follow-up, 18 from 60-week follow-up, 16 from 2-year follow-up and 6 from 3-year follow-up. Forty-eight of the 142 pairs of ratings (33.8%) were made at a point at which the respective participant was classed to be a 'case' (i.e., have an eating disorder of clinical severity), whereas 94 (66.2%) of the pairs of ratings where from time points when the respective participant was no longer judged to have a clinical eating disorder. Since more than one set of data was used per participant, statistical independence could be questioned. Therefore a variety of multilevel models, with a random effect for subject, were fitted to the data to allow for correlation between repeated measures (Singer & Willett, 2003).

The relationship between the Global Impairment Index and the clinicians' ratings was assessed using a two-tailed bivariate Spearman's correlation with alpha set at 0.01. This was done in two ways: first, all available data (142 pairs of ratings) were entered into the analysis and Spearman's correlation coefficient was calculated; and second, the data on both

measures were compared at each individual assessment time point to test whether the agreement between the two measures was independent of the time point of assessment.

In addition, a one-way ANOVA was performed to examine the relationship between the Global Impairment Index and the severity of the clinicians' ratings on the EDE-based 0 to 6 point scale. Since a clinicians' rating of 6 (the highest rating) was made only once over the 142 occasions, this category was combined with ratings of 5. Post hoc tests for comparisons between each rating were performed using Tukey's B procedure based on an alpha coefficient of 0.01. Finally, a test of linearity was performed.

***Second test of construct validity**: Relationship between the CIA-Q Global Impairment Index and the global EDE-Q score*

The dataset for this sub-study was constructed by using CIA-Q and EDE-Q data of each of the 123 participants only once (i.e., from one assessment time point). Therefore, each participant 'contributed' one pair of ratings only to the analysis. Participants' 'first-time' CIA-Q data (i.e., the CIA-Q data from the assessment time at which a participant first completed the CIA-Q) and the EDE-Q data from that respective time point were chosen for analysis[21]. The data distribution was as follows: data from the beginning of treatment were available for 48 participants, from the end of treatment for 32, from 20-week follow-up for 9, from 40-week follow-up for 20, from 60-week follow-up for 10, and from 2-year follow-up for four participants. Sixty-nine of the 123 participants (56.1%) were classified by one of the three expert clinicians as having an eating disorder of clinical severity at the time of assessment, whereas 54 (43.9%) no longer had a clinical eating disorder.

The relationship between Global Impairment Index and global EDE-Q score was assessed using a two-tailed bivariate Spearman's correlation with alpha set at 0.01.

[21] This was done in order to ensure that no participant had completed the CIA-Q more than once.

Test of discriminant validity

The dataset for this sub-study was constructed by using CIA-Q data and eating disorder caseness information of each of the 127 participants only once (i.e., from one assessment time point). Participants' 'first-time' CIA-Q data and the eating disorder case status at the respective assessment time were chosen for analysis. The data distribution for this study was as follows: data from the beginning of treatment were used for 50 participants, from the end of treatment for 32, from 20-week follow-up for 9, from 40-week follow-up for 20, from 60-week follow-up for 11, and from 2-year follow-up for 5 participants. Seventy-two of the 127 participants (56.7%) were judged to have a clinical eating disorder at the time that they completed the CIA-Q, the others no longer having an eating disorder of clinical severity.

An independent samples t-test was performed in order to evaluate whether participants who no longer had an eating disorder scored significantly lower on the CIA-Q than participants who were classified as still having a clinical eating disorder.

Results

First test of construct validity: *Relationship between the CIA-Q Global Impairment Index and the clinicians' impairment ratings*

The mean and standard deviation for the Global Impairment Indices and clinicians' ratings, and the correlation coefficient obtained for them are shown in Table 6.5.

Table 6.5 *First test of construct validity: Spearman's correlation for CIA-Q GII[1] and clinicians' impairment rating at each assessment time and in total*

Assessment time	N	CIA-Q GII[1] M (SD)	Clinicians' rating M (SD)	r (p)
Beginning of treatment	26	1.94 (0.69)	4.31 (0.74)	.644 (.000)**
End of treatment	26	0.47 (0.34)	2.12 (1.18)	.402 (.042)*
20-week follow-up	23	0.59 (0.48)	1.78 (1.24)	.414 (.049)*
40-week follow-up	27	0.46 (0.37)	1.70 (1.30)	.472 (.013)*

60-week follow-up	18	0.56 (0.51)	1.50 (1.38)	.618 (.006)**
2-year follow-up	16	0.49 (0.46)	1.69 (1.35)	.612 (.012)*
3-year follow-up	6	0.79 (0.74)	2.67 (1.75)	.924 (.008)**
Total	142	0.79 (0.74)	2.28 (1.56)	.704 (.000)**

[1] Global Impairment Index * significant on 0.05 level ** significant on 0.01 level

The table shows the correlation coefficient for each time point and for all pairs of ratings (total). There was a statistically significant positive correlation between the two measures at every assessment point. This was also true for the total sample of measurements. Thus, a higher CIA-Q Global Impairment Index was associated with a higher rating of psychosocial impairment as judged by the clinicians.

The one-way ANOVA revealed that there was a statistically significant difference in Global Impairment Index between the severity ratings (0 to 5) made by the clinicians ($F(5,0)=54.5$, $p<0.001$). The mean Global Impairment Index for each individual clinical rating is shown in Table 6.6. Tukey post hoc tests showed that there were no statistically significant differences between clinicians' ratings of 0, 1 and 2 with regard to the mean Global Impairment Indices of participants given these ratings. However, clinicians' ratings of 3, 4 and 5 each differed significantly from every other rating (i.e., the mean Global Impairment Index of participants who received a clinical rating of 3, 4 or 5 was significantly different from that of participants with every other rating).

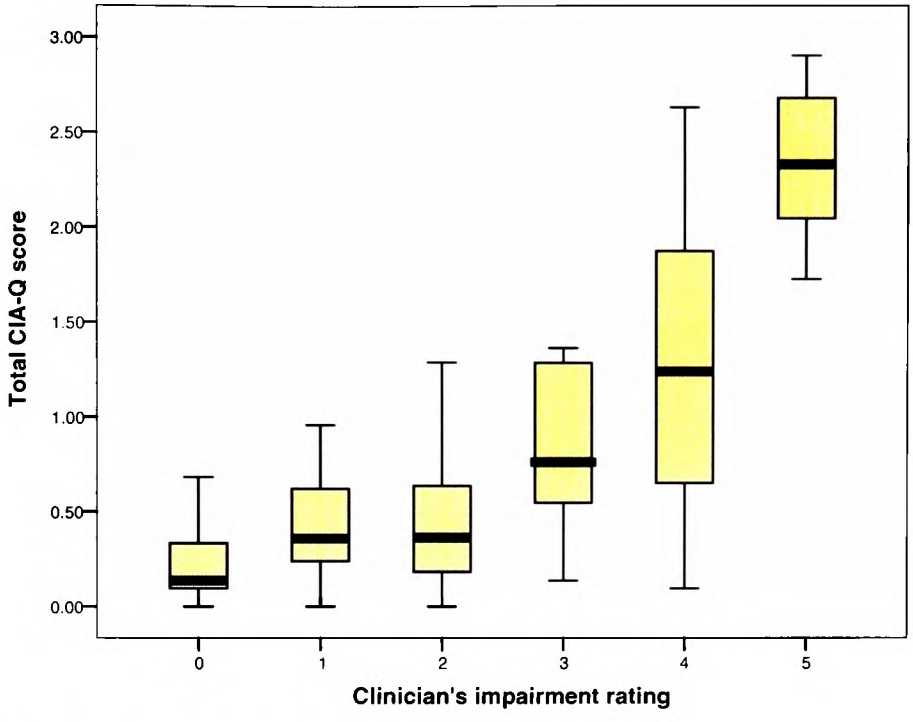
Table 6.6 *First test of construct validity: Mean CIA-Q GII[1] for each clinicians' rating group (0 to 5)*

Clinician's rating	N	CIA-Q GII[1] M (SD)
0	25	0.22 (0.20)
1	18	0.43 (0.33)
2	41	0.47 (0.37)
3	21	0.86 (0.38)*
4	24	1.29 (0.74)*
5	13	2.32 (0.37)*
Total	142	0.79 (0.74)

[1] Global Impairment Index
* Statistically significantly different to any other clinician's rating group; p=.05

Figure 6.4 shows a boxplot diagram of the distribution of the mean Global Impairment Index per rating. It can be seen that participants who received a rating of 5 (the highest impairment rating) scored highest on the CIA-Q, and that lower impairment ratings were associated with lower scores on the CIA-Q. Participants who received a rating of 0 ('no impairment') scored the lowest. While there was a linear trend between Global Impairment Index and clinicians' impairment ratings ($F(1)=225.4$, $p<0.001$) (indicating that CIA-Q scores increased linearly with increase in clinicians' ratings), there was also a significant deviation from linearity ($F(4)=11.8$, $p<0.001$). This can be explained by the fact that the mean Global Impairment Index for the rating groups 0, 1 and 2 did not differ significantly, whereas the differences between the Global Impairment Indices at the higher end of the clinical impairment rating (ratings 3, 4 and 5 particularly) were much larger (see Figure 6.4).

Figure 6.4 *First test of construct validity: Boxplot of mean CIA-Q GII[1] in the five clinicians' rating groups (0 to 5)*

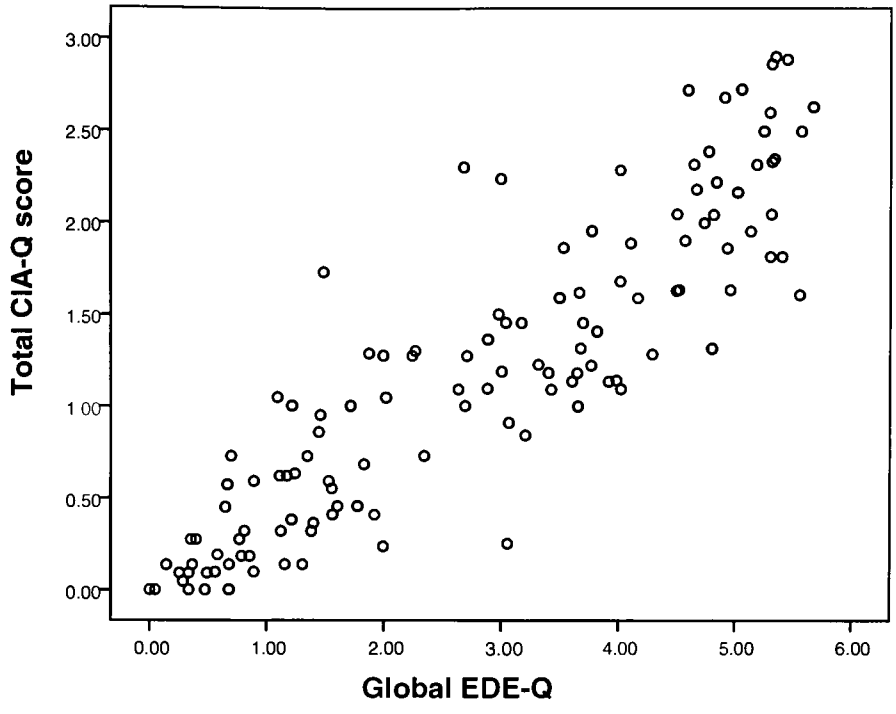


[1] Global Impairment Index

Second test of construct validity: *Relationship between the CIA-Q Global Impairment Index and the global EDE-Q score*

The mean Global Impairment Index for the 123 participants was 1.19 (SD=0.83) and the mean global EDE-Q score was 2.81 (SD=1.74). There was a statistically significant positive correlation (r_s=0.90) between the CIA-Q score and global EDE-Q score, indicating that higher levels of secondary psychosocial impairment were associated with higher levels of eating disorder psychopathology. Figure 6.5 shows a scatter plot of the scores on the two measures.

Figure 6.5 *Second test of construct validity: Scatter plot showing the relationship between CIA-Q GII[1] and global EDE-Q score*



[1] Global Impairment Index

Test of discriminant validity

The mean Global Impairment Index of the 55 participants who were judged to no longer have an eating disorder of clinical severity (M=0.50, SD=0.43) was lower than that of the 72 participants who were classified as still having an eating disorder (M=1.74, SD=0.64). This difference was statistically significant (t (123.12)=-12.97, p<0.001).

Summary of findings

- It was found that there was a strong positive correlation between the ratings of impairment made by clinicians and the degree of impairment rated by patients on the CIA-Q. Higher clinicians' ratings were associated with higher CIA-Q Global Impairment Indices.

- A one-way ANOVA revealed that there was a statistically significant difference in Global Impairment Index between the six severity ratings of impairment made by the clinicians. The mean Global Impairment Indices of participants who received a clinician's rating of either 3 or 4 or 5 were significantly different from those of participants with any other rating.
- There was a strong positive correlation between Global Impairment Index and the severity of accompanying eating disorder psychopathology as measured by the EDE-Q.
- Participants who were judged to no longer have a clinical eating disorder had significantly lower CIA-Q scores than those who were judged to still have an eating disorder.
- These findings suggest that the CIA-Q is a valid measure of the severity of psychosocial impairment secondary to eating disorder psychopathology as assessed and rated by an experienced clinician. The CIA-Q also appears to possess good discriminant validity.

6.3.4 Study Four: Sensitivity to Change of the CIA-Q

Design

The aim of this study was to evaluate the sensitivity to change of the CIA-Q. This was tested in two ways. First, pre- and post-treatment CIA-Q data were compared of patients receiving a new transdiagnostic form of cognitive behaviour therapy for eating disorders (Fairburn et al., 2003), were compared. Most patients with eating disorders respond well to this treatment and show a substantial decrease in eating disorder psychopathology (Fairburn, 2004; 2005). If the CIA-Q is sensitive to change in secondary

psychosocial impairment, then participants should score significantly lower on the instrument after they have received this treatment.

The second test was a more sensitive version of the first. Change in patients' secondary psychosocial impairment from the beginning to the end of treatment, as judged by change in CIA-Q Global Impairment Index, was compared to change in psychosocial impairment as rated by one of two expert eating disorder clinicians (Dr Zafra Cooper or Professor Christopher Fairburn). It was predicted that, if the instrument was sensitive to change, there would be a positive association between the degree of change detected by the clinician and the extent of the change in CIA-Q score.

Participants

As in the previous studies, the participants for this study were selected from the 170 patients with a clinical eating disorder described in Chapter Three. For the first test of sensitivity to change, the CIA-Q data of 36 of the 170 patients were used. These were the patients for whom CIA-Q data were available at both the beginning and end of treatment.

The second test of sensitivity to change was only conducted in Oxford. The data of 21 of the 105 Oxford patients were used. These were patients for whom data on the CIA-Q and a clinician's rating were available at the beginning and end of treatment.

Measures

The two measures for this study, the CIA-Q and the clinician's ratings of secondary psychosocial impairment have been described in Sections 6.2 and 6.3.3 of this chapter respectively.

Data analysis

In order to determine whether participants' total score on the CIA-Q was significantly lower after treatment than before, a paired samples t-test was performed. In

order to examine whether the degree of change in psychosocial impairment after treatment (compared to before) as detected by an expert clinician was positively associated with that detected by the CIA-Q, change scores for both the Global Impairment Index and the clinicians' ratings were calculated, indicating the difference in ratings between the beginning and end of treatment. A Spearman's correlation coefficient for these change scores was calculated with alpha set at 0.01.

Results

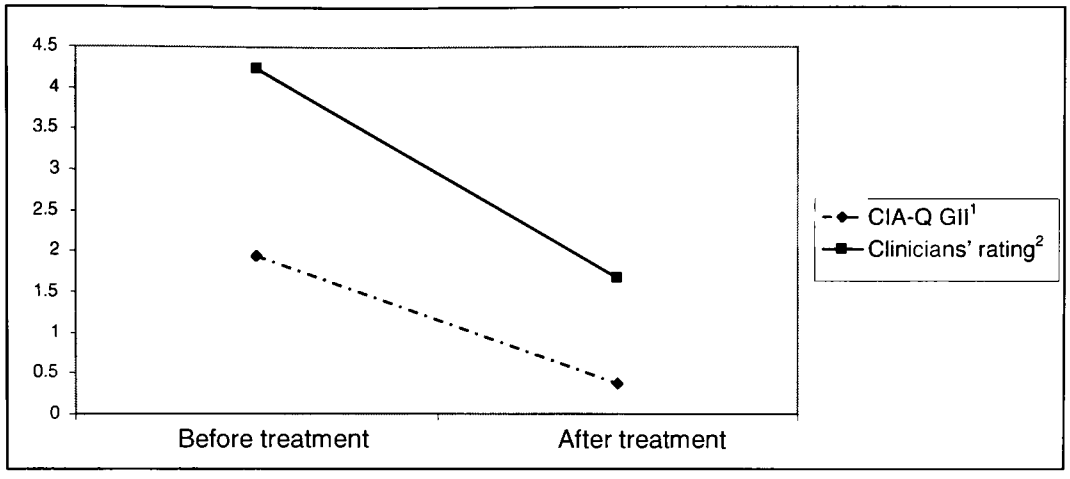
***Test 1**: Change after treatment with cognitive behaviour therapy*

There was a significant difference between participants' mean Global Impairment Indices at the beginning and end of treatment ($t(35)=12.99$, $p<0.001$). Participants' mean score dropped from 1.83 (SD=0.71) at the start of treatment to 0.45 (SD=0.48) at the end of treatment.

***Test 2**: Comparison of change on the CIA-Q and clinician-rated change*

The mean Global Impairment Index of the 21 participants was 1.93 (SD=0.74) at the beginning of treatment and 0.38 (SD=0.38) at the end, the mean change score being 1.55 (SD=0.64). The pre-treatment mean for the clinicians' rating of secondary psychosocial impairment was 4.24 (SD=0.70) and 1.67 (SD=1.11) at the end of treatment, the mean change score being 2.57 (SD=1.33). There was a significant positive correlation between the two change scores ($r_s=.553$, $p=0.009$). Figure 6.6 shows the severity of secondary psychosocial impairment before and after treatment as assessed by the CIA-Q (dashed line) and the clinicians' rating (solid line).

Figure 6.6 *Change in CIA-Q GII[1] and clinician's rating[2] of impairment in 21 patients before and after treatment.*



[1] Global Impairment Index; Range: 0 to 3
[2] Range: 0 to 6

Summary of findings

- Patients' CIA-Q Global Impairment Indices were significantly lower at the end of 20 weeks of cognitive behaviour therapy than at the beginning.
- There was a significant positive association between the change in secondary psychosocial impairment (from beginning to end of treatment) as rated on the CIA-Q and the change rated by an expert clinician.

6.4 Discussion

The overall aim of this set of studies was to establish the psychometric properties of the CIA-Q and test its reliability and validity. Four studies were conducted. The findings of each study are summarised in turn and then discussed jointly. The chapter ends with conclusions drawn from all for studies and suggestions for further research.

Study One

The aim of the first study was to test the internal consistency and the factorial structure of the instrument. The CIA-Q showed excellent internal consistency with a

Cronbach's Alpha of .96. All of the 22 items of the CIA-Q showed significant positive correlations with the total CIA-Q score, ranging from .48 to .86. The principal component analysis revealed the presence of three components with eigenvalues higher than 1. After inspection of the scree plot it was decided to retain two components for the performance of Oblimin rotation. It was found that 18 of the 22 items loaded highly on the first component and four items loaded highly on the second component. All of these latter items were questions about secondary impairment in work performance and had an 'if applicable' qualifier, meaning that participants should leave these items blank if they were not currently working. Twenty-two percent of the 127 participants left the work impairment items blank. This is likely to have accounted for these items grouping together on a second component. Also, as mentioned earlier, the scree plot showed a clear break after the first component and only a minor one after the second, suggesting that most items contributed only to the first. This suggests that the CIA-Q measures just one construct. Given the content of the instrument and the fact that the subsequent studies showed that the CIA-Q scores were closely related to secondary psychosocial impairment as judged by an expert clinician, the construct measured by the CIA-Q is likely to be psychosocial impairment secondary to eating disorder psychopathology.

Study Two

The aim of the second study was to evaluate the test-retest reliability of the CIA-Q. This involved administering the CIA-Q twice within three days to a group of subjects with varying degrees of eating disorder psychopathology. It was found that there was a statistically significant positive correlation between participants' CIA-Q Global Impairment Indices at the two time points. The Global Impairment Index was slightly lower at time point two than at time point one, but this difference was small. This suggests that the CIA-Q

has satisfactory test-retest reliability. The finding that the Global Impairment Index on the second administration of the CIA-Q was slightly lower than that on the first is consistent with regression to the mean.

Study Three

The aim of the third study was to test the construct and the discriminant validity of the CIA-Q. Construct validity was examined in two ways. First, the performance of the CIA-Q was compared to the best available method for assessing psychosocial impairment as a result of eating disorder psychopathology, the judgment of an experienced clinician. The second test was an indirect one. It was based on the assumption that there is a positive association between the severity of eating disorder psychopathology and the severity of the psychosocial impairment secondary to it. Scores on the CIA-Q were compared with the severity of eating disorder psychopathology present as assessed using the EDE-Q. It was predicted that, if the CIA-Q is a valid measure of psychosocial impairment due to eating disorder features, a clinician's judgment about the severity of impairment a patient is experiencing and their global EDE-Q score should both be positively correlated with the total CIA-Q score, the Global Impairment Index.

The discriminant validity of the CIA-Q was tested by comparing the total CIA-Q score of patients with an eating disorder of clinical severity with a group of subjects who were no longer judged to have a clinical eating disorder. It was predicted that, if the CIA-Q had satisfactory discriminant validity, the CIA-Q scores of participants with a clinical eating disorder should be significantly higher than those of participants without an eating disorder.

It was found that there was a strong positive correlation between the ratings of impairment made by a clinician and the degree of impairment rated by patients on the CIA-Q. A higher Global Impairment Index was associated with a higher clinicians' rating. A

one-way ANOVA revealed that there was a statistically significant difference in CIA-Q score between the six-point severity ratings made by the clinician. The mean Global Impairment Index of participants who received a clinician's rating of either 3 or 4 or 5 was significantly different from that of participants with any other rating. The CIA did not distinguish between participants with low levels of impairment as rated by one of the two clinicians; that is, the Global Impairment Indices of participants receiving a clinicians' impairment rating of 0, 1 or 2 did not differ significantly.

With regard to the second test of construct validity it was found that there was a statistically significant positive correlation between Global Impairment Index and global EDE-Q score, indicating that higher levels of secondary psychosocial impairment were associated with higher levels of eating disorder psychopathology.

As regards the test of discriminant validity, there was a statistically significant difference in total CIA-Q score between participants who had an eating disorder of clinical severity and those who were judged to no longer have a clinical eating disorder.

The findings of these three studies all support the validity of the CIA-Q. The instrument showed good construct validity by matching the judgment of two expert eating disorder clinicians concerning the severity of secondary psychosocial impairment. It also correlated highly with the severity of associated eating disorder psychopathology. Finally, the CIA-Q successfully discriminated between participants with and without a clinical eating disorder.

Study Four

The aim of the fourth study was to test the sensitivity to change of the CIA-Q. This was done in two ways. First, it was assessed whether there was a significant decrease in patients' Global Impairment Indices after they had received treatment for their eating

disorder. Second, change in participants' secondary psychosocial impairment from the beginning of treatment to the end, as judged by the change in their Global Impairment Indices, was compared to the change in impairment as rated by two expert eating disorder clinicians. It was found that patients' CIA-Q scores were significantly lower at the end of treatment than at the beginning. Furthermore, it was found that there was a significant positive association between the change on the CIA-Q and change in psychosocial impairment as rated by an expert clinician.

The findings across both tests of sensitivity to change suggest that the CIA-Q is able to detect change in the severity of psychosocial impairment secondary to eating disorder features.

Methodological considerations

Common to all four studies were the following strengths. The overall sample from which the participants were selected (see Chapter Three) was well characterised. Leading measures of eating disorder psychopathology were employed. Given the two-centre catchment area sampling frame, the patients are likely to have been representative of many other outpatient samples of adults with eating disorders. A further strength is that access to this sample enabled the assessment of the same participants before, and at several time points after, the treatment for their eating disorder. Therefore, CIA-Q data were available from participants covering the whole range of severity of eating disorder features (i.e., from mild to severe). This was particularly important given the goal of testing the validity and the performance of the CIA-Q, which required testing the whole scoring range of the instrument.

In addition, a strength of Study Two was the comparison of the performance of the CIA-Q with a detailed fine-grain assessment of the severity of secondary psychosocial impairment by an expert eating disorder clinician, the latter being a particularly strong

validator for a self-report instrument of the same construct. The judgment of an expert clinician is currently the only way of assessing impairment secondary to eating disorder features.

The main limitation of the studies is the relatively small sample size, especially in Study Four. Data subsets from the total sample of the 170 patients described in Chapter Three were used. The reasons for doing so have been described. This reduced the amount of available data, thereby reducing statistical power. Nevertheless, significant associations were found, highlighting the validity of the CIA-Q, which was evident even with this small sample size. Another limitation is that the findings of the studies are based on a single sample. Furthermore the sample contained hardly any patients with anorexia nervosa. Replication of the findings with other, preferably larger, samples containing more patients with anorexia nervosa is necessary to generalise from them. Another limitation is that the sample was a British one. It is unknown whether the findings can be generalised to samples elsewhere, and samples with a different ethnic mix in particular. A further limitation is that the clinicians' ratings of the severity of psychosocial impairment were made by just two expert clinicians. It is unknown whether the judgments of these two clinicians are representative of the judgments of other expert eating disorder clinicians. Another potential limitation is the fact that participants' CIA-Q data from several time points were included in the analyses. This could have compromised statistical independence. However, the data were modelled to remove any confounding time or person effects (Singer & Willett, 2003), and therefore the use of repeated measures should not have influenced the results. A further limitation of Study Two is that no CIA-Q data of healthy controls were acquired. The discriminant validity of the CIA-Q was established using a group of patients with a clinical eating disorder and a group of people who no longer had an eating disorder. This was a

strong test of the discriminant validity of the CIA-Q. Nevertheless it would be desirable to acquire normative data for the CIA-Q (i.e., the impairment that stems from normative levels of eating disorder features). Finally, a potential limitation of the instrument is that there is no assessment of impairment of physical functioning due to eating disorder features. As explained above, it was a deliberate decision to exclude physical functioning, since it was thought that it would be impossible for patients with an eating disorder to rate the full impact their eating disorder is having on their physical health. This said, information on the likely physical health of a patient can be gathered from the EDE-Q, which has been chosen as the partner instrument of the CIA-Q (e.g., body mass index, frequency of self-induced vomiting and laxative misuse).

Conclusions

This chapter describes the development and validation of a brief self-report measure of the severity of psychosocial impairment secondary to eating disorder psychopathology, the Clinical Impairment Assessment - Questionnaire (CIA-Q). No such measure currently exists. The findings show that the CIA-Q has sound psychometric properties, is a valid measure of the severity of psychosocial impairment secondary to eating disorder psychopathology and is sensitive to change.

CHAPTER SEVEN
Conclusions and Implications

7.1 Chapter overview

This chapter is concerned with the broader significance of the work described in this dissertation and with potential future lines of research. In Section 7.2 the overall aims of the dissertation are restated and the findings summarised. In Section 7.3 the main conclusions are described. In Section 7.4 the broader relevance of the work is discussed, and in Section 7.5 future directions for research are proposed.

7.2 Restatement of aims and summary of findings

The research described in this dissertation had four overarching aims. The first was to describe the clinical characteristics of a large and representative sample of patients with eating disorder NOS and compare them with those of patients with bulimia nervosa. The second was to examine how the problems of nosology and neglect associated with this DSM-IV diagnosis might be solved. Three solutions were proposed. The third aim was to derive a provisional operational definition of what constitutes an "eating disorder". This involved developing an interview-based measure of functional impairment secondary to eating disorder features and administering it to a large sample of people exhibiting the full range of eating disorder psychopathology. The fourth aim was to develop a clinically useful measure of psychosocial impairment secondary to eating disorder features.

With regard to the first aim it was found that patients with eating disorder NOS were remarkably similar to those with bulimia nervosa with regard to their demographic features, the full range of their psychopathology and the duration of their eating disorder. Few differences were found, and these were to be expected given the definition of the two diagnoses.

With regard to the second aim it was found that applying the first of the three suggested solutions (viz. systematically relaxing the diagnostic criteria for anorexia nervosa and bulimia nervosa along lines suggested in the literature) to a representative sample of outpatients with a clinical eating disorder proved not to be sufficient to reduce the high prevalence of eating disorder NOS to a "residual" level. Eating disorder NOS remained the most common eating disorder diagnosis. The second of the proposed solution was therefore relevant (viz. relaxing the diagnostic criteria for anorexia nervosa and bulimia nervosa as above, and then re-classifying the remaining cases). It was found that a group of expert eating disorder clinicians judged this second solution to be an improvement over the existing scheme for classifying eating disorders and an acceptable one in the short term. Interestingly, a substantial number of these clinicians thought that the third and most radical of the three solutions (viz. the "transdiagnostic" solution in which no eating disorder subtypes were recognised) was the best if a long-term perspective were to be taken.

With regard to the third aim, that is to derive a provisional operational definition of what constitutes an eating disorder, specific severity levels on five eating disorder features were identified that were strongly associated with the presence of a clinically significant level of impairment. These eating disorder features were: the pursuit of strict dietary rules, objective bulimic episodes, purging episodes, dissatisfaction with shape and weight, and over-concern with maintaining strict control over eating. The presence of two or more of these features above the identified threshold was most predictive of a clinically significant level of impairment. Thus, an impairment-based, transdiagnostic, provisional operational definition of an eating disorder of clinical severity was derived.

With regard to the fourth aim, an easily-administered self-report measure of the severity of psychosocial impairment secondary to eating disorder features was developed.

Studies of its psychometric properties, reliability, validity and sensitivity to change all supported its use.

The strengths and limitations of each of these studies have been considered in each of the respective chapters.

7.3 Conclusions of the research

A number of general conclusions can be drawn from the present findings. First, the studies of eating disorder NOS showed that eating disorder NOS is not only common but it is also severe. The findings indicate that it is comparable in severity to the established eating disorder bulimia nervosa. This needs to be taken into account in any revision to the DSM or ICD systems for classifying eating disorders. In an article written jointly with Professor Fairburn, the candidate proposed two ways that this could be done with respect to the DSM-IV scheme. One involved re-casting the eating disorder NOS cases into two groups, either binge eating disorder or a new eating disorder diagnosis (perhaps termed mixed eating disorder) after subthreshold presentations of anorexia nervosa and bulimia nervosa had been subsumed by these two diagnoses. This strategy accepts the current classificatory framework and the notion of sub-classifying eating disorders into separate categories. Were this solution to be implemented it would bring to attention the large and neglected group of eating disorder NOS cases and would be useful in the short-term. The second proposed solution was more radical and involved abandoning all diagnostic distinctions within the eating disorders and examining their classification afresh (the transdiagnostic solution) so that potential new, empirically-based, subdivisions could be derived. Both solutions raise a major issue; the absence of an agreed definition of what is an eating disorder. Such a definition is needed to define eating disorder NOS (as it has no positive diagnostic criteria) and to implement the transdiagnostic solution. In line with the definition of a psychiatric

disorder, defining what is an eating disorder of clinical severity requires identifying those eating disorder features that lead to a clinically significant level of impairment. Doing so is currently limited by the absence of a measure of impairment secondary to eating disorder features. The candidate has developed such a measure and used it to derive, for the first time, an operational definition of an eating disorder. Finally, the research has highlighted that the topic of secondary impairment is not only neglected in eating disorder research, but also in clinical practice. To help direct attention at impairment, an easily-administered self-report measure of secondary psychosocial impairment was developed.

7.4 Broader relevance of the research

The current research not only has implications for the American Psychiatric Association's classificatory system for eating disorders (DSM-IV; APA, 1994) but also for the one provided in the International Classification of Diseases (ICD-10) of the World Health Organisation (WHO, REF). The ICD-10 classificatory system recognises eight different eating disorders. These are as follows: 1) anorexia nervosa, 2) atypical anorexia nervosa, 3) bulimia nervosa, 4) atypical bulimia nervosa, 5) overeating associated with other psychological disturbances, 6) vomiting associated with other psychological disturbances, 7) other eating disorders, 8) eating disorder, unspecified. The boundaries between these different eating disorders are vague and the categories overlap to a degree. How these eating disorder subtypes relate to the presentations seen in clinical practice is not known nor is there any empirical justification for this plethora of eating disorder diagnoses. As for DSM-IV, but possibly to a greater extent, there is a need for ICD to use an empirically-based classificatory scheme.

What about eating disorders in children and adolescents? The research in this dissertation has been exclusively concerned with the diagnosis and classification of eating

disorders in adults. One cannot assume that the same findings and conclusions would apply to children and adolescents with eating disorders where the clinical presentations can be quite different (Bryant-Waugh, 2000). This said, it has been pointed out that, as with adults, the DSM-IV scheme for classifying eating disorders fits children and adolescents poorly and that the most common diagnostic group is eating disorder NOS (Nicholls, Chater & Lask, 2000).

Finally, the research has implications for other psychiatric conditions in that the strategy used for defining what is an eating disorder could be used for defining the outer boundaries of other psychiatric disorders.

7.5 Future directions

First and foremost, the need for replication of the current research needs to be stressed. It would be particularly desirable for there to be further studies of the reliability and validity of both the CIA and the CIA-Q using larger and more ethnically diverse samples and ones containing more patients with anorexia nervosa. The use of different validators for the two instruments (i.e. different expert clinicians, other measures of impairment) would also be desirable. If the reliability and validity of the CIA were to be replicated, an important next step would be to use the instrument with different samples of eating disorder patients in order to see if the operational definition of an eating disorder derived in this dissertation was wholly or partially replicated. If so, the definition could be used in many different ways; for example, in epidemiological studies of the prevalence of eating disorders and in treatment studies as a major index of outcome. It could also be utilised to examine the classification of eating disorders afresh from the transdiagnostic perspective.

Another important line of future research would be to study further functional impairment secondary to eating disorder features. Both the CIA and the CIA-Q could be used to study the relationship between impairment and specific eating disorder features. Knowledge of this type could be used to inform the development of therapeutic interventions designed to reduce impairment secondary to treatment-resistant symptoms. The relationship between the performance of the CIA and the CIA-Q could also be usefully examined, much as the relationship between the EDE and EDE-Q has been researched (e.g., Fairburn & Beglin, 1994).

Lastly, and as mentioned above, the strategy for defining what is an eating disorder could be applied to other psychiatric disorders, and the outer boundaries of these conditions could be derived.

REFERENCES

Abraham, S. F., Brown, T., Boyd, C., Luscombe, G., & Russell, J. (2006). Quality of life: eating disorders. *Australian and New Zealand Journal of Psychiatry, 40,* 150-155.

Abraham, S. F. & Lovell, N. (1999). Research and clinical assessment of eating and exercise behaviour. *Hospital Medicine, 60,* 481-485.

Abraham, S. F., Pettigrew, B., Boyd, C., Russell, J., & Taylor, A. (2005). Usefulness of amenorrhoea in the diagnoses of eating disorder patients. *Journal of Psychosomatic Obstetrics and Gynecology, 26,* 211-215.

Agras, W. S., Crow, S. J., Halmi, K. A., Mitchell, J. E., Wilson, G. T., & Kraemer, H. C. (2000a). Outcome predictors for the cognitive behavior treatment of bulimia nervosa: data from a multisite study. *American Journal of Psychiatry, 157,* 1302-1308.

Agras, W. S., Walsh, B. T., Fairburn, C. G., Wilson, G. T., & Kraemer, H. C. (2000b). A multicenter comparison of cognitive-behavioral therapy and interpersonal psychotherapy for bulimia nervosa. *Archives of General Psychiatry, 57,* 459-466.

American Psychiatric Association (1980). *DSM-III: Diagnostic and Statistical Manual of Mental Disorders.* (3 ed.) Washington,D.C.: American Psychiatric Association.

American Psychiatric Association (1987). *DSM-III-R.* Washington,D.C.: American Psychiatric Association.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition).* Washington, D.C.: American Psychiatric Association.

Austoker, J. (1994). Reducing alcohol intake. *British Medical Journal, 308,* 1549-1552.

Bartlett, M. S. (1954). A note on the multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society, 16,* 296-298.

Beaumont, P. J. V., Kopec-Schrader, E. M., Talbot, P., & Touyz, S. W. (1993). Measuring the specific psychopathology of eating disorder patients. *Australian and New Zealand Journal of Psychiatry, 27,* 506-511.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. E., & Erbaugh, J. K. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4,* 561-571.

Beglin, S. J. (1990). *Eating Disorders in Young Adult Women.* Oxford: Doctor of Philosophy thesis, University of Oxford.

Ben-Tovim, D. I. (1988). DSM-III, draft DSM-III-R, and the diagnosis and prevalence of bulimia in Australia. *American Journal of Psychiatry, 145,* 1000-1002.

Black, C. M. D. & Wilson, G. T. (1996). Assessment of eating disorders: Interview versus questionnaire. *International Journal of Eating Disorders, 20,* 43-50.

Blaskovich, J. & Tomaka, J. (1993). Measures of self-esteem. In J.P.Robinson, P. R. SHAVER, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 115-160). San Diego: Academic Press.

Boini, S., Briancon, S., Guillemin, F., Galan, P., & Hercberg, S. (2004). Impact of cancer occurrence on health-related quality of life. *Health and Quality of Life Outcomes, 9,* 4.

Bryant-Waugh, R. (2000). Overview of the eating disorders. In B.Lask & R. Bryant-Waugh (Eds.), *Anorexia Nervosa and Related Eating Disorders in Childhood and Adolescence* (2 ed., pp. 27-40). Hove: Psychology Press.

Bulik, C. M., Sullivan, P. F., Fear, J., & Pickering, A. (1997). Predictors of the development of bulimia nervosa in women with anorexia nervosa. *Journal of Nervous and Mental Disease, 185,* 704-707.

Bulik, C. M., Sullivan, P. F., & Kendler, K. S. (2000). An empirical study of the classification of eating disorders. *American Journal of Psychiatry, 157,* 886-895.

Button, E. J., Benson, E., Nollett, C., & Palmer, R. L. (2005). Don't forget EDNOS (eating disorder not otherwise specified): patterns of service use in an eating disorders service. *Psychiatric Bulletin, 29,* 134-136.

Cachelin, F. M. & Maher, B. A. (1998). Is amenorrhea a critical criterion for anorexia nervosa? *Journal of Psychosomatic Research, 44,* 435-440.

Catell, R. B. (1966). The scree test for number of factors. *Multivariate Behavioral Research, 1,* 245-276.

Clinton, D., Button, E., Norring, C., & Palmer, R. (2004). Cluster analysis of key diagnostic variables from two independent samples of eating-disorder patients: evidence for a consistent pattern. *Psychological Medicine, 34,* 1035-1045.

Clinton, D. & Norring, C. (1999). The Rating of Anorexia and Bulimia (RAB) interview: Development and preliminary validation. *European Eating Disorders Review, 7,* 362-371.

Clinton, D. & Norring, C. (2005). The comparative utility of statistically derived eating disorder clusters and DSM-IV diagnoses: Relationship to symptomatology and psychiatric comorbidity at intake and follow-up. *Eating Behaviors, 6,* 403-418.

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

Cooper, P. J., Osborn, M., Gath, D., & Feggetter, G. (1982). Evaluation of a modified self-report measure of social adjustment. *British Journal of Psychiatry, 141,* 68-75.

Cooper, Z., Cooper, P. J., & Fairburn, C. G. (1989). The validity of the Eating Disorder Examination and its subscales. *British Journal of Psychiatry, 154,* 807-812.

Cooper, Z. & Fairburn, C. (1987). The Eating Disorder Examination: a semistructured interview for the assessment of the specific psychopathology of eating disorders. *International Journal of Eating Disorders, 6,* 1-8.

Cooper, Z. & Fairburn, C. G. (2003). Refining the definition of binge eating disorder and nonpurging bulimia nervosa. *International Journal of Eating Disorders, 34,* S89-S95.

Cotrufo, P., Barretta, V., Monteleone, P., & Maj, M. (1998). Full-syndrome, partial-syndrome and subclinical eating disorders: An epidemiological study of female students in Southern Italy. *Acta Psychiatrica Scandinavica, 98,* 112-115.

Cotrufo, P., Monteleone, P., Castaldo, E., & Maj, M. (2004). A 4-year epidemiological study of typical and atypical eating disorders: Preliminary evidence for subgroups of atypical eating disorders with different natural outcomes. *European Eating Disorders Review, 12,* 234-239.

Couturier, J. L. & Lock, J. (2006). Do supplementary items on the eating disorder examination improve the assessment of adolescents with anorexia nervosa? *International Journal of Eating Disorders, 39,* 426-433.

Crow, S. J., Agras, W. S., Halmi, K., Mitchell, J. E., & Kraemer, H. C. (2002). Full syndromal versus subthreshold anorexia nervosa, bulimia nervosa, and binge eating disorder: A multicenter study. *International Journal of Eating Disorders, 32,* 309-318.

De la Rie, S. M., Noordenbos, G., & van Furth, E. F. (2005). Quality of life and eating disorders. *Quality of Life Research, 14,* 1511-1522.

Derogatis, L. R. (1975). *Brief Symptom Inventory.* Baltimore: Clinical Psychosomatic Research.

Derogatis, L. R. (1977). *The SCL-90 Manual: Scoring, Administration and Procedures for the SCL-90.* Baltimore: Clinical Psychometric Research.

Derogatis, L. R. & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine, 13,* 595-605.

Doll, H. A., Petersen, S. E., & Stewart-Brown, S. L. (2005). Eating disorders and emotional and physical well-being: Associations between student self-reports of eating disorders and quality of life as measured by the SF-36. *Quality of Life Research, 14,* 705-717.

Engel, S. G., Wittrock, D. A., Crosby, R. D., Wonderlich, S. A., Mitchell, J. E., & Kolotkin, R. L. (2006). Development and psychometric validation of an eating dilsorder-specific health-related quality of life instrument. *International Journal of Eating Disorders,* 62-71.

Fairburn, C. G. (2005). Evidence-based treatment of eating disorders: problems and solutions. *Journal of Psychosomatic Research, 58,* S2.

Fairburn, C. G., Agras, W. S., Walsh, B. T., Wilson, G. T., & Stice, E. (2004). Prediction of outcome in bulimia nervosa by early change in treatment. *American Journal of Psychiatry, 161,* 2322-2324.

Fairburn, C. G. & Beglin, S. J. (1994). Assessment of eating disorders: interview or self-report questionnaire? *International Journal of Eating Disorders, 16,* 363-370.

Fairburn, C. G. & Bohn, K. (2005). Eating disorder NOS (EDNOS): an example of the troublesome "not otherwise specified" (NOS) category in DSM-IV. *Behaviour Research and Therapy, 43,* 691-701.

Fairburn, C. G. & Cooper, P. J. (1984). The clinical features of bulimia nervosa. *British Journal of Psychiatry, 144,* 238-246.

Fairburn, C. G. & Cooper, Z. (1993). The Eating Disorder Examination (12th edition). In C.G.Fairburn & G. T. Wilson (Eds.), *Binge Eating: Nature, Assessment and Treatment* (pp. 317-360). New York: Guilford Press.

Fairburn, C. G., Cooper, Z., Doll, H. A., Norman, P., & O'Connor, M. (2000). The natural course of bulimia nervosa and binge eating disorder in young women. *Archives of General Psychiatry, 57,* 659-665.

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: a "transdiagnostic" theory and treatment. *Behaviour Research and Therapy, 41,* 509-528.

Fairburn, C. G. & Gowers, S. G. (in press). Eating disorders. In M.Rutter, D. Bishop, D. Pine, S. Scott, J. Stevenson, E. Taylor, & A. Thapar (Eds.), *Rutter's Child and Adolescent Psychiatry* (5th ed.), Oxford: Blackwell.

Fairburn, C. G. & Harrison, P. J. (2003). Eating disorders. *Lancet, 361,* 407-416.

Fairburn, C. G. & Walsh, B. T. (2002). Atypical eating disorders (eating disorder not otherwise specified). In C.G.Fairburn & K. D. Brownell (Eds.), *Eating disorders and obesity: a comprehensive handbook* (2nd ed., pp. 171-177). New York: Guilford Press.

Fairburn, C. G., Welch, S. L., & Hay, P. J. (1993). The classification of recurrent overeating: The "binge eating disorder" proposal. *International Journal of Eating Disorders, 13,* 155-159.

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

Garfinkel, P. E. & Garner, D. M. (1982). *Anorexia Nervosa: A Multidimensional Perspective.* New York: Brunner/Mazel.

Garfinkel, P. E., Kennedy, S. H., & Kaplan, A. S. (1995a). Views on classification and diagnosis of eating disorders. *Canadian Journal of Psychiatry, 40,* 445-456.

Garfinkel, P. E., Lin, E., Goering, P., Spegg, C., Goldbloom, D., Kennedy, S. et al. (1996). Should amenorrhoea be necessary for the diagnosis of anorexia nervosa? Evidence from a Canadian community sample. *British Journal of Psychiatry, 168,* 500-506.

Garfinkel, P. E., Lin, E., Goering, P., Spegg, C., Goldbloom, D. S., Kennedy, S. et al. (1995b). Bulimia nervosa in a Canadian community sample: Prevalence and comparison of subgroups. *American Journal of Psychiatry, 152,* 1052-1058.

Garner, D. M. (1995). Assessment of eating disorder psychopathology. In K.D.Brownell & C. G. Fairburn (Eds.), *Eating Disorders and Obesity.* (pp. 117-121). New York: Guilford Press.

Garner, D. M., Olmsted, M. P., & Polivy, J. (1983). Development and validation of a mutidimensional Eating Disorder Inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders, 2,* 15-34.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4,* 26-42.

Grilo, C. M. (2006). *Eating and weight disorders.* New York: Psychology Press.

Grilo, C. M., Masheb, R. M., Lozano-Blanco, C., & Barry, D. T. (2004). Reliability of the Eating Disorder Examination in patients with binge eating disorder. *International Journal of Eating Disorders, 35,* 80-85.

Hathaway, S. R. & McKinley, J. C. (1943). *The Minnesota Multiphasic Personality Inventory.* Minneapolis: University of Minnesota Press.

Hay, P. (2003). Quality of life and bulimic eating disorder behaviors: Findings from a community-based sample. *International Journal of Eating Disorders, 33,* 434-442.

Hay, P. J., Fairburn, C. G., & Doll, H. A. (1996). The classification of bulimic eating disorders: A community-based cluster analysis study. *Psychological Medicine, 26,* 801-812.

Hebebrand, J., Casper, R., Treasure, J., & Schweiger, U. (2004). The need to revise the diagnostic criteria for anorexia nervosa. *Journal of Neural Transmission, 111,* 827-840.

Herzog, D. B., Hopkins, J. D., & Burns, C. D. (1993). A follow-up study of 33 subdiagnostic eating disordered women. *International Journal of Eating Disorders, 14,* 261-267.

Herzog, D. B., Norman, D. K., Rigotti, N. A., & Pepose, M. (1986). Frequency of bulimic behaviors and associated social maladjustment in female graduate students. *Journal of Psychiatric Research, 20,* 355-361.

Johnson, C. L., Stuckey, M. K., Lewis, L. D., & Schwartz, D. M. (1983). A survey of 509 cases of self-reported bulimia. In P.L.Darby, P. E. Garfinkel, D. M. Garner, & D. V. Coscina (Eds.), *Anorexia Nervosa: Recent Developments in Research* (pp. 159-171). New York: Alan R. Liss.

Kaiser, H. (1974). An index of factorial simplicity. *Psychometrika, 39,* 31-36.

Keel, P. K., Fichter, M., Quadflieg, N., Bulik, C. M., Baxter, M. G., Thornton, L. et al. (2004). Application of a latent class analysis to empirically define eating disorder phenotypes. *Archives of General Psychiatry, 61,* 192-200.

Keel, P. K., Haedt, A., & Edler, C. (2005). Purging disorder: An ominous variant of bulimia nervosa? *International Journal of Eating Disorders, 38,* 191-199.

Keel, P. K., Mayer, S. A., & Harnden-Fischer, J. H. (2001). Importance of size in defining binge eating episodes in bulimia nervosa. *International Journal of Eating Disorders, 29,* 294-301.

Kendler, K. S. (1999). Setting boundaries for psychiatric disorders. *American Journal of Psychiatry, 156,* 1845-1848.

Kolotkin, R. L., Crosby, R. D., Kosloski, K. D., & Williams, G. R. (2001). Development of a brief measure to assess quality of life in obesity. *Obesity Research, 9,* 102-111.

Kolotkin, R. L., Head, S., Hamilton, M., & Tse, C. K. (1995). Assessing impact of weight on quality of life. *Obesity Research, 3,* 49-56.

Kolotkin, R. L., Westman, E. C., Ostbye, T., Crosby, R. D., Eisenson, H. J., & Binks, M. (2004). Does binge eating disorder impact weight-related quality of life? *Obesity Research, 12,* 999-1005.

Kushner, R. F. & Foster, G. D. (2000). Obesity and quality of life. *Nutrition, 16,* 947-952.

Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159-174.

Las Hayas, C., Quintana, J. M., Padierna, A., Bilbao, A., Munoz, P., Madrazo, A. et al. (2006). The new questionnaire Health-Related Quality of Life for Eating Disorders showed good validity and reliability. *Journal of Clinical Epidemiology, 59,* 192-200.

Le Grande, D., Binford, R. B., Peterson, C. B., Crow, S. J., Crosby, R. D., Klein, M. H. et al. (2006). DSM-IV threshold versus subthrehold bulimia nervosa. *International Journal of Eating Disorders, 39,* 462-467.

Lee, S., Ho, T. P., & Hsu, L. K. G. (1993). Fat phobic and nonfat phobic anorexia nervosa - A comparative study of 70 chinese patients in Hong Kong. *Psychological Medicine, 23,* 999-1017.

Loeb, K. L., Pike, K. M., Walsh, B. T., & Wilson, G. T. (1994). Assessment of diagnostic features of bulimia nervosa: Interview versus self-report format. *International Journal of Eating Disorders, 16,* 75-81.

Luce, K. H. & Crowther, J. H. (1999). The reliability of the Eating Disorder Examination-self-report questionnaire version (EDE-Q). *International Journal of Eating Disorders, 25,* 349-351.

Malgady, R. G., Rogler, L. H., & Tryon, W. W. (1992). Issues of validity in the Diagnostic Interview Schedule. *Journal of Psychiatric Research, 26,* 59-67.

Mannucci, E., Ricca, V., Di Bernardo, M., Moretti, S., Cabras, P. L., & Rotella, C. M. (1998). Psychometric properties of EDE 12.0D in obese adult patients without binge eating disorder. *Eating and Weight Disorders, 2,* 144-149.

Mannucci, E., Ricca, V., Di Bernardo, M., & Rotella, C. M. (1996). Studio del comportamento alimentare con una intervista strutturata: La Eating Disorder Examination. *Il Diabete, 8,* 127-131.

Marcus, M. D., Wing, R. R., & Fairburn, C. G. (1995). Cognitive treatment of binge eating v. behavioral weight control in the treatment of binge eating disorder. *Annals of Behavioral Medicine, 17,* SO90.

Martin, C. K., Williamson, D. A., & Thaw, J. M. (2000). Criterion validity of the multiaxial assessment of eating disorders symptoms. *International Journal of Eating Disorders, 28,* 303-310.

Masheb, R. M. & Grilo, C. M. (2004). Quality of life in patients with binge eating disorder. *Eating and Weight Disorders, 9,* 194-199.

Milos, G., Spindler, A., Schnyder, U., & Fairburn, C. G. (2005). Instability of eating disorder diagnoses: prospective study. *British Journal of Psychiatry, 187,* 573-578.

Mitchell, J. E., Cook-Myers, T., & Wonderlich, S. A. (2005). Diagnostic criteria for anorexia nervosa: Looking ahead to DSM- V. *International Journal of Eating Disorders, 37,* S95-S97.

Mitchell, J. E., Pyle, R. L., Hatsukami, D., & Eckert, E. D. (1986). What are atypical eating disorders? *Psychosomatics, 27,* 21-28.

Mizes, J. S. & Sloan, D. M. (1998). An empirical analysis of eating disorder, not otherwise specified: Preliminary support for a distinct subgroup. *International Journal of Eating Disorders, 23,* 233-242.

Mond, J. M., Hay, P. J., Rodgers, B., & Owen, C. (2006). An update on the definition of "excessive exercise" in eating disorders research. *International Journal of Eating Disorders, 39,* 147-153.

Mond, J. M., Hay, P. J., Rodgers, B., Owen, C., & Beumont, P. J. V. (2004a). Relationships between exercise behaviour, eating-disordered behaviour and quality of life in a community sample of women: When is exercise 'excessive'? *European Eating Disorders Review, 12,* 265-272.

Mond, J. M., Hay, P. J., Rodgers, B., Owen, C., & Beumont, P. J. V. (2005). Assessing quality of life in eating disorder patients. *Quality of Life Research, 14,* 171-178.

Mond, J. M., Rodgers, B., Hay, P., Korten, A., Owen, C., & Beumont, P. (2004b). Disability associated with community cases of commonly occurring eating disorders. *Australian and New Zealand Journal of Public Health, 28,* 246-251.

National Collaborating Centre for Mental Health (2004). *Eating Disorders: Core Interventions in the Treatment and Management of Anorexia Nervosa, Bulimia Nervosa and Related Eating Disorders.* London: British Psychological Society and Royal College of Psychiatrists.

Nevonen, L. & Broberg, A. G. (2005). A comparison of sequenced individual and group psychotherapy for eating disorder not otherwise specified. *European Eating Disorders Review, 13,* 29-37.

Nicholls, D., Chater, R., & Lask, B. (2000). Children into DSM don't go: A comparison of classification systems for eating disorders in childhood and early adolescence. *International Journal of Eating Disorders, 28,* 317-324.

Nielsen, S. & Palmer, B. (2003). Diagnosing eating disorders - AN, BN and the others. *Acta Psychiatrica Scandinavica, 108,* 161-162.

Nollett, C. L. & Button, E. J. (2005). Questionnaire measures of psychopathology in eating disorders: Comparisons between clinical groups. *European Eating Disorders Review, 13,* 211-215.

Norvell, N. & Cooley, E. (1986). Diagnostic issues in eating disorders - Two cases of atypical eating disorder. *International Journal of Psychiatry in Medicine, 16,* 317-323.

Padierna, A., Quintana, J. M., Arostegui, I., Gonzalez, N., & Horcajo, M. O. J. (2000). The health-related quality of life in eating disorders. *Quality of Life Research, 9,* 667-674.

Pallant, J. (2005). *SPSS survival manual.* (2 ed.) New York: Open University Press.

Palmer, B. (2003). Concepts of eating disorders. In J.Treasure, U. Schmidt, & E. van Furth (Eds.), *Handbook of eating disorders* (pp. 1-10). Chichester: Wiley.

Palmer, R., Christie, M., Cordle, C., Davies, D., & Kenrick, J. (1987). The Clinical Eating Disorder Rating Instrument (CEDRI) - A preliminary description. *International Journal of Eating Disorders, 6,* 9-16.

Palmer, R. L. (1993). Weight concern should not be a necessary criterion for the eating disorders: a polemic. *International Journal of Eating Disorders, 14,* 459-465.

Pincus, H. A., Wakefield Davis, W., & McQueen, L. E. (1999). "Subthreshold" mental disorders: A review and synthesis of studies on minor depression and other "brand names". *British Journal of Psychiatry, 174,* 288-296.

Pratt, E. M., Niego, S. H., & Agras, W. S. (1998). Does the size of a binge matter? *International Journal of Eating Disorders, 24,* 307-312.

Ramacciotti, C. E., Dell'Osso, L., Paoli, R. A., Ciapparelli, A., Coli, E., Kaplan, A. S. et al. (2002). Characteristics of eating disorder patients without a drive for thinness. *International Journal of Eating Disorders, 32,* 206-212.

Ricca, V., Mannucci, E., Mezzani, B., Di Bernardo, M., Zucchi, T., Paionni, A. et al. (2001). Psychopathological and clinical features of outpatients with an eating disorder not otherwise specified. *Eating and Weight Disorders, 6,* 157-165.

Rieger, E., Touyz, S. W., Swain, T., & Beumont, P. J. V. (2001). Cross-cultural research on anorexia nervosa: Assumptions regarding the role of body weight. *International Journal of Eating Disorders, 29,* 205-215.

Rieger, E., Wilfley, D. E., Stein, R. I., Marino, V., & Crow, S. J. (2005). A comparison of quality of life in obese individuals with and without binge eating disorder. *International Journal of Eating Disorders, 37,* 234-240.

Rizvi, S. L., Peterson, C. B., Crow, S. J., & Agras, W. S. (2000). Test-retest reliability of the eating disorder examination. *International Journal of Eating Disorders, 28,* 311-316.

Rosen, J. C., Vara, L., Wendt, S., & Leitenberg, H. (1990). Validity studies of the Eating Disorder Examination. *International Journal of Eating Disorders, 9,* 519-528.

Rosenberg, M. (1965). *Society and the Adolescent Self-Image.* Princeton: Princeton University 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.

Russell, G. F. M. (1969). Metabolic, endocrine and psychiatric aspects of anorexia nervosa. *The Scientific Basis of Medicine Annual Reviews, 15,* 236-255.

Serreti, A., Olgiati, P., & Colombo, C. (2005). Components of self-esteem in affective patients and non-psychiatric controls. *Journal of Affective Disorders, 88,* 93-98.

Shisslak, C. M., Crago, M., & Yates, A. (1989). Typical patterns in atypical anorexia nervosa. *Psychosomatics, 30,* 307-311.

Singer, J. D. & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* New York: Oxford University Press.

Sloan, D. M., Mizes, J. S., & Epstein, E. M. (2005). Empirical classification of eating disorders. *Eating Behaviors, 6,* 53-62.

Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory STAI (form Y) and the State-Trait Anxiety Inventory for children STAIC (form C).* Palo Alto: Consulting Psychologist Press, Inc.

Spilker, B. & Revicki, R. (1996). Quality of life. In B.Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (pp. 185-197). Philadelphia: Lippincott-Raven.

Spitzer, R. L., Stunkard, A., Yanovski, S., Marcus, M. D., Wadden, T., Wing, R. et al. (1993). Binge eating disorder should be included in DSM-IV A reply to Fairburn et al's "The classification of recurrent overeating - The binge eating disorder proposal". *International Journal of Eating Disorders, 13,* 161-169.

Spitzer, R. L., Williams, J. B. W., Gibbon, M., & First, M. B. (1992). The structured clinical interview for DSM-III-R (SCID): I. History, rationale and description. *Archives of General Psychiatry, 49,* 624-629.

Steiger, H. & Ghadirian, A. M. (1989). Atypical eating disorders resembling anorexia nervosa - A report of five cases. *International Journal of Eating Disorders, 8,* 307-314.

Strober, M., Freeman, R., & Morrell, W. (1999). Atypical anorexia nervosa: Separation from typical cases in course and outcome in a long-term prospective study. *International Journal of Eating Disorders, 25,* 135-142.

Stunkard, A. J. & Allison, K. C. (2003). Binge eating disorder - Disorder or marker? *International Journal of Eating Disorders, 34,* S107-S116.

Sullivan, P. F., Bulik, C. M., Carter, F. A., Gendall, K. A., & Joyce, P. R. (1996). The significance of a prior history of anorexia in bulimia nervosa. *International Journal of Eating Disorders, 20,* 253-261.

Sysko, R., Walsh, B. T., & Fairburn, C. G. (2005). Eating Disorder Examination-Questionnaire as a measure of change in patients with bulimia nervosa. *International Journal of Eating Disorders, 37,* 100-106.

Tabachnick, B. G. & Fidell, L. S. (2001). *Using multivariate statistics.* (4 ed.) New York: Harper Collins.

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 and Weight Disorders, 6,* 121-129.

Thurstone, L. L. (1947). *Multiple factor analysis.* Chicago: University of Chicago Press.

Treasure, J. (1999). *Anorexia nervosa: A survival guide for families, friends and sufferers.* Psychology Press.

Turner, H. & Bryant-Waugh, R. (2004). Eating disorder not otherwise specified (EDNOS): profiles of clients presenting at a community eating disorder service. *European Eating Disorders Review, 12,* 18-26.

Waller, G. (1993). Why do we diagnose different types of eating disorder? Arguments for a change in research and clinical practice. *Eating Disorders Review, 1,* 74-89.

Walsh, B. T. (1993). Binge eating in bulimia nervosa. In C.G.Fairburn & G. T. Wilson (Eds.), *Binge Eating: Nature, Assessment and Treatment.* (pp. 37-49). New York: Guilford Press.

Walsh, B. T., Fairburn, C. G., Mickley, D., Sysko, R., & Parides, M. K. (2004). Treatment of bulimia nervosa in a primary care setting. *American Journal of Psychiatry, 161,* 556-561.

Walsh, B. T. & Garner, D. M. (1997). Diagnostic issues. In D.M.Garner & P. E. Garfinkel (Eds.), *Handbook of treatment for eating disorders.* (2 ed., pp. 25-33). New York: Guilford Press.

Ware, J. E., Kosinski, M., & Keller, S. D. (1996). A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care, 34,* 220-233.

Ware, J. E., Snow, K. K., Kosinski, M., & Reese P.R. (1993). *SF-36 health survey: Manual and interpretation guide.* Boston, Mass: New England Medical Center.

Watson, T. L. & Andersen, A. E. (2003). A critical examination of the amenorrhea and weight criteria for diagnosing anorexia nervosa. *Acta Psychiatrica Scandinavica, 108,* 175-182.

Weissman, M. M. & Bothwell, S. (1976). Assessment of social adjustment by patient self-report. *Archives of General Psychiatry, 33,* 1111-1115.

Weissman, M. M., Olfson, M., Gameroff, M. J., Feder, A., & Fuentes, M. (2001). A comparison of three scales for assessing social functioning in primary care. *American Journal of Psychiatry, 158,* 460-466.

Weissman, M. M. & Paykel, E. S. (1974). *The depressed woman: A study of social relationships.* Chicago: University of Chicago Press.

Weissman, M. M., Prusoff, B. A., Thompson, W. D., Harding, P. S., & Myers, J. K. (1978). Social adjustment by self-report in a community sample and in psychiatric out-patients. *Journal of Nervous and Mental Disease, 166,* 317-326.

WHOQOL Group (1998a). Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychological Medicine, 28,* 551-558.

WHOQOL Group (1998b). The World Health Organization quality of life assessment (WHOQOL): Development and general psychometric properties. *Social Science and Medicine, 46,* 1569-1585.

Wilfley, D. E. & Cohen, L. R. (1997). Psychological treatment of bulimia nervosa and binge eating disorder. *Psychopharmacology Bulletin, 33,* 437-454.

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., Welch, R. R., Stein, R. I., Spurrell, E. B., Cohen, L. R., Saelens, B. E. et al. (2002). A randomized comparison of group cognitive-behavioral therapy and group

interpersonal psychotherapy for the treatment of overweight individuals with binge eating disorder. *Arch.Gen.Psychiatry, 59,* 713-721.

Wilfley, D. E., Wilson, G. T., & Agras, W. S. (2003). The clinical significance of binge eating disorder. *International Journal of Eating Disorders, 34,* S96-S106.

Williamson, D. A., Gleaves, D. H., & Savin, S. S. (1992). Empirical classification of eating disorder not otherwise specified: Support for DSM-IV changes. *Journal of Psychopathology and Behavioral Assessment, 14,* 201-216.

Wilson, G. T. & Eldredge, K. L. (1991). Frequency of binge eating in bulimic patients: diagnostic validity. *International Journal of Eating Disorders, 10,* 557-561.

Wilson, G. T. & Smith, D. (1989). Assessment of bulimia nervosa: An evaluation of the Eating Disorder Examination. *International Journal of Eating Disorders, 8,* 173-179.

Woodside, D. B., Garfinkel, P. E., Lin, E., Goering, P., Kaplan, A. S., Goldbloom, D. S. et al. (2001). Comparisons of men with full or partial eating disorders, men without eating disorders, and women with eating disorders in the community. *American Journal of Psychiatry, 158,* 570-574.

World Health Organization (1992). *The ICD-10 Classification of Mental and Behavioural Disorders - Clinical Descriptions and Diagnostic Guidelines.* Geneva: WHO.

World Health Organization (2002). *Reducing risks, promoting healthy life - The World Health Report.* Geneva.

INDEX OF TABLES

Table 1.1.....13
Table 1.2.....14
Table 1.3.....16
Table 1.4.....17
Table 1.5.....22
Table 1.6.....29
Table 3.1.....67
Table 3.2.....68
Table 3.3.....75
Table 3.4.....78
Table 3.5.....86
Table 3.6.....89
Table 3.7.....93
Table 4.1.....114
Table 4.2.....121
Table 4.3.....124
Table 4.4.....124
Table 4.5.....126
Table 4.6.....127
Table 5.1.....144
Table 5.2.....146
Table 5.3.....147
Table 5.4.....155
Table 5.5.....159
Table 5.6.....163
Table 5.7.....166
Table 5.8.....169
Table 5.9.....171
Table 5.10.....172
Table 5.11.....173
Table 5.12.....174

Table 5.13.....178
Table 5.14.....181
Table 5.15.....188
Table 5.16.....192
Table 5.17.....194
Table 5.18.....196
Table 5.19.....198
Table 5.20.....200
Table 5.21.....201
Table 6.1.....217
Table 6.2.....224
Table 6.3.....225
Table 6.4.....227
Table 6.5.....237
Table 6.6.....239

INDEX OF FIGURES

Figure 1.1.....30
Figure 3.1.....72
Figure 3.2.....77
Figure 3.3.....80
Figure 3.4.....88
Figure 3.5.....92
Figure 4.1.....116
Figure 5.1.....148
Figure 5.2.....176
Figure 5.3.....197
Figure 6.1.....226
Figure 6.2.....230
Figure 6.3.....234
Figure 6.4.....240
Figure 6.5.....241
Figure 6.6.....245

INDEX OF APPENDICES

Appendix 1.1 Paper by Fairburn & Bohn (2005), published in *Behaviour Research and Therapy*, regarding the problems associated with eating disorder NOS **274**

Appendix 3.1 Eating Disorder Examination 15.0 (EDE; Fairburn & Cooper, 1993) **285**

Appendix 3.2 Brief Symptom Inventory (BSI; Derogatis & Melisaratos, 1983) **333**

Appendix 3.3 Social Adjustment Scale – Modified version (SAS-M; Cooper et al., 1982) **335**

Appendix 3.4 Rosenberg Self-Esteem Scale (RSE; Rosenberg, 1965) **342**

Appendix 4.1 Self-report Questionnaire sent to the registrants of the 2004 annual meeting of the Eating Disorder Research Society in Amsterdam **343**

Appendix 5.1 Clinical Impairment Assessment – Interview (CIA; Version 3.0) **346**

Appendix 5.2 Scoring Sheet for the CIA **357**

Appendix 5.3 EDE items excluded from further analysis for conceptual reasons **361**

Appendix 6.1 Clinical Impairment Assessment – Questionnaire (CIA-Q version 1.0) **362**

Appendix 6.2 Eating Disorder Examination – Questionnaire (EDE-Q; Fairburn & Beglin, 1994) **364**

Appendix 6.3 Re-test version of the CIA-Q **367**