Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Eating Behaviors 6 (2005) 403–418
EATING BEHAVIORS

The comparative utility of statistically derived eating disorder clusters and DSM-IV diagnoses: Relationship to symptomatology and psychiatric comorbidity at intake and follow-up

David Clinton [a,*], Claes Norring [b,c]

[a]Division of Psychiatry, M57, Neurotec Department, Karolinska Institutet, Huddinge University Hospital, S-141 86, Sweden
[b]Psychiatric Research Centre, Örebro County Council, P.O. Box 1613, S-701 16 Örebro, Sweden
[c]Department of Behavioral, Social and Legal Sciences (Psychology), Örebro University, Sweden

Received 21 January 2005; received in revised form 4 May 2005; accepted 26 May 2005

Abstract

Introduction: The classification of eating disorders has been a matter of considerable debate. The present paper extends previous work and aimed to compare the utility of statistically derived clusters of eating disorders and conventional diagnoses.

Methods: Adult female eating disorder patients who had previously been classified on the basis of cluster analysis of key diagnostic variables were examined on measures of eating disorder symptomatology and psychiatric comorbidity at intake ($N=601$) and subsequent follow-up after 6 and 36 months ($N=349$, $N=322$, respectively).

Results: Compared to DSM-IV diagnoses, clusters demonstrated greater utility in terms of more distinct between-group differences and higher effect sizes in relation to a wide range of variables. The greater utility of clusters was in important respects due to the reallocation of EDNOS patients to more relevant alternative categories and to a greater emphasis on psychological and behavioural features of eating disorders.

Conclusions: In order to achieve a better classification of eating disorders, it will be important to place increased emphasis on common psychological features. There is a need to move away from increased use of subtypes and toward a definition of eating disorder per se.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Anorexia nervosa; Bulimia nervosa; Classification; Comorbidity; Diagnosis; Eating disorders

* Corresponding author.
E-mail address: David.Clinton@neurotec.ki.se (D. Clinton).

1471-0153/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eatbeh.2005.05.003

1. Introduction

The classification of eating disorders is unsatisfactory in several important respects. Prevailing diagnostic criteria for these disorders used within the DSM (APA, 1994) and ICD (WHO, 1992) systems tend to be based on clinical opinion and consensus. As such, they tend to be arbitrary and do not necessarily reflect empirically derived truths about the utility of these criteria. The DSM system, which is most often used by both researchers and clinicians, has undergone considerable revision during the past quarter century. Revisions of criteria relating to the eating disorders have largely focused on achieving increased precision pertaining to the two major eating disorder syndromes, namely, anorexia nervosa (AN) and bulimia nervosa (BN). This has meant becoming more specific about weight (AN) and defining thresholds in terms of frequency and duration for criteria relating to amenorrhoea (AN), binge eating (BN) and compensatory behaviour after binges (BN).

For example, DSM-III specified weight loss of at least 25% of original body weight as a criterion for AN, which meant that in extreme cases, even patients who were overweight but had experienced rapid and extreme weight loss could fulfil this criterion. DSM-IV has been amended to specify 15% of expected weight, which has led to the establishment of cut-off points, usually computed in terms of body mass index (BMI). Since the advent of DSM-III-R amenorrhoea for at least three consecutive menstrual cycles has been required for a diagnosis of AN among females. As regards BN, which was not described until the seminal article of Russell (1979), and which first entered the DSM system the following year as bulimia, a number of specifications have occurred. Among other criteria, DSM-III-R specified episodes of binge eating that occurred at least twice weekly over the past 3 months along with regular use of compensatory behaviour (e.g., vomiting, use of laxatives, excessive exercise, fasting) to limit weight gain. In DSM-IV, this latter criterion was amended in a similar fashion to the binge-eating criterion and now specifies compensatory behaviour that occurs at least twice weekly over the past 3 months. Another addition to DSM-IV was the criterion that "self-evaluation is unduly influenced by shape and weight". This criterion was added after much controversy in order to emphasise what many researchers saw as an essential similarity between BN and AN.

At best, a system of classification should provide categories that are mutually exclusive and collectively exhaustive. However, the present DSM system largely fails these tests (Palmer, 2003). Categories overlap or are prevented from doing so only by arbitrary rules, and individuals may move from one diagnosis to another over time. Perhaps the most unsatisfactory aspect of the current DSM-IV diagnostic system is its reliance on the residual category of eating disorder not otherwise specified (EDNOS). Patients consigned to this sizable "rag bag" category have clear eating problems but do not fulfil criteria for the two major syndromes. The diagnosis may, for example, cover patients of extremely low weight who fail to meet the amenorrhoea criterion for AN, or overweight patients who binge and use compensatory behaviour, but not sufficiently frequently to fulfil criteria for BN. According to recent studies (Fairburn & Harrison, 2003; Turner & Bryant-Waugh, 2004), roughly half of all patients seeking help with an eating disorder do not fulfil criteria for AN or BN and must therefore be assigned to the EDNOS category.

Empirical attempts at clarifying the classification of eating disorders have focused on the use of factor analysis and cluster analysis and have been reviewed in our previous work (Clinton, Button, Norring, & Palmer, 2004). Both techniques have been used for many years by psychiatric researchers investigating problems of diagnosis (Everitt & Landau, 1998). While factor analysis reveals patterns among variables, cluster analysis focuses on groupings of individuals. Although these studies provide some useful

empirical indicators about the classification of eating disorders, they also have important shortcomings. They have tended to focus on specific diagnoses (Hay, Fairburn, & Doll, 1996) or sub-groups of diagnoses (Grilo, Masheb, & Berman, 2001; Mizes & Sloan, 1998; Stice & Agras, 1999) rather than the entire spectrum of eating disorders. Samples have often been small (Grilo et al., 2001; Mizes & Sloan, 1998; Van der Ham, Meulman, van Strien, & van Engeland, 1997) or based on non-clinical cases with doubtful relevance to clinical eating disorders (Bulik, Sullivan, & Kendler, 2000). The latter study utilised latent class analysis as opposed to cluster analysis in their taxometric approach. This method was also recently used by Keel et al. (2004) who found that a four-class solution supported some of the distinctions drawn by DSM-IV, while introducing new features to better define eating disorder phenotypes.

Our approach to the question of how eating disorders can best be classified has been based on cluster analysis (Everitt, Landau, & Leese, 2001). In previous work by the present authors (Clinton et al., 2004), we used cluster analysis to explore natural groupings of patients who present a wide range of eating disorder services, using data on key diagnostic variables routinely collected on a series of patients presented to 15 different centres in Sweden and to 1 centre in England. Our results suggested that the classification of eating disorders in both samples could be approached most fruitfully from the standpoint of three distinct clusters of patients. The largest cluster in both samples 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 of "anorexics" was characterised by low weight, amenorrhoea and the absence of binge eating and seemed to correspond to the clinical picture of restricting anorexia nervosa. The third cluster of "overeaters" was characterised by high weight and moderate levels of binge eating and compensatory behaviour. It was concluded that the patterns from both countries resembled, but were not identical to, existing diagnostic categories.

Our findings are worth comparing with the recent work of Sloan, Mizes, and Epstein (2005), who also utilised cluster analysis to explore the classification of eating disorders among 159 patients from five distinct clinics. Instead of using key diagnostic variables as we did in our previous paper, the Sloan group clustered patients on the eight EDI subscales along with data on weight, binge eating, and purging. Although the Sloan study suffered a number of statistical shortcomings (i.e., no attempt was made to exclude outliers prior to cluster analysis; no statistical criteria were used to help determine the optimal number of clusters; and no use was made of non-hierarchical cluster techniques), their results are not dissimilar to our own. They identified four clusters of patients and concluded that there is a relatively poor fit between empirically derived groupings of patients and clinical diagnoses.

Although previous work by others and ourselves sheds light on questions pertaining to the natural grouping of patients with eating disorders, these studies say nothing about the comparative utility of statistically derived clusters in relation to conventional diagnoses. This is an important issue since if we are to consider alternative schemes of classification, any such alternatives must be shown to have greater practical utility. Utility can be explored by examining the relationship of clusters and diagnoses to relevant measures of eating disorder symptoms and comorbidity at initial assessment and over time. If one of the systems demonstrates greater differentiation of categories, along with an ability to explain a higher percentage of variation, at intake and follow-up, then we have evidence of the comparative utility of that particular scheme of classification. The present study attempts to address this issue and expands our previous work using the data collected in Sweden, which offered both a wider variety of comorbidty measures and follow-up data than the English data did. Primary aims of the study were to explore the

clinical utility of previously defined statistically derived clusters of eating disorder patients by comparing clusters and DSM-IV diagnostic categories on important aspects of comorbidity assessed initially at intake and subsequently after 6 and 36 months.

2. Methods

2.1. Participants

Adult female eating disorder patients from Sweden who were studied in our previous work participated in the present study ($N=601$). The sample was collected within the framework of the Co-ordinated Evaluation and Research at Specialist Units for Eating Disorders in Sweden (CO-RED) Project, a longitudinal naturalistic study of the treatment of eating disorders at 15 specialist centres. Units offer a wide variety of treatment forms such as inpatient, day patient, outpatient, individual psychotherapy, family and group therapy, psychoactive drugs, expressive forms of treatment, etc. The distribution of DSM-IV diagnoses was AN ($N=137$, 22.8%), BN ($N=240$, 39.9%), BED ($N=31$, 5.2%), and EDNOS, ($N=193$, 32.1%). All subjects provided informed consent to take part in the CO-RED study. Age ranged from 14 to 49 years ($M=24.5$ years, S.D. = 6.4). Mean duration of eating disorder at presentation was 8.2 years (S.D. = 6.7). Data were subsequently obtained on the same measures after 6 months ($N=349$) and after 36 months ($N=322$). Since only 58% of patients could be followed up after 6 months and 54% after 36 months, dropout analysis was conducted. When patients who could be assessed at follow-up were compared with dropouts on the measures used in the study using one-way ANOVA, both groups were virtually indistinguishable with only a few exceptions. Patients who attended 6-month follow-up had previously reported significantly less (i.e., $p<.05$) fear of weight gain and laxative abuse at initial assessment. At 36-month follow-up, patients who could be assessed had previously reported significantly greater disturbed body image and had scored significantly higher on EDI Asceticism. The distribution of initial DSM-IV diagnoses at 6-month follow-up sample was AN ($N=86$, 24.6%), BN ($N=132$, 37.8%), BED ($N=24$, 6.9%), and EDNOS ($N=107$, 30.7%). Among the patients assessed after 36 months, the distribution of initial DSM-IV was AN ($N=75$, 23.3%), BN ($N=126$, 39.1%), BED ($N=16$, 5.0%), and EDNOS ($N=105$, 32.6%).

2.2. Instruments

The Rating of Anorexia and Bulimia Interview (RAB) was used to assess key diagnostic variables (Clinton & Norring, 1999; Nevonen, Broberg, Clinton, & Noring, 2003). The RAB is a 56-item semi-structured interview with graded response formats covering a wide range of eating disorder symptoms, concomittant psychopathology and background variables; it generates operational DSM-IV eating disorder diagnoses and is widely used in Sweden. It has satisfactory internal consistency and inter-rater reliability; kappa ranged from .47 to .92 ($M=.74$) for the variables used in the present study (Nevonen et al., 2003). From the RAB, 10 essential clinical variables for the diagnosis of eating disorders according to DSM-IV were selected for subsequent cluster analysis. These 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, amenorrhea, and body image disturbance. The RAB was also used to obtain interview-based measures of comorbidity pertaining to drug and

alcohol problems, suicidal behaviour and self-harm, as well as relationship problems. Follow-up versions of the RAB involved asking the same questions as were asked at intake, but in relation to the patient's present state at the time. In the case of the 6-month follow-up, a shortened version of the RAB was used with questions only relating to eating disorder symptoms. This was done in order to shorten the time staff and participants needed to complete assessments. Full interviews were conducted at 36 months. Further assessment of eating disorder symptoms was made using the self-report questionnaire Eating Disorders Inventory-2 (EDI-2) (Garner, Olmsted, & Polivy, 1983). Psychiatric symptoms were measured using a shortened (63-item) version of the self-report questionnaire Symptom Check List-90 (SCL-90) (Derogatis, Lipman, & Covi, 1973). The SCL was shortened by removing the subscales for Phobic Anxiety, Paranoid Ideation, Psychoticism and Additional Scales.

2.3. Procedure

Data were collected by staff from participating units. Interviewers had long experience in the assessment of eating disorders in a clinical setting using the respective instruments. For the most part, interviewers were either qualified psychiatrists or clinical psychologists, although other professionals, such as experienced nurses and social workers, also took part. Training of interviewers took place at participating units. Centrally arranged project meetings and workshops were also used for training of interviewers and for making checks on how instruments were being used once the project was underway. Administration of interviews and self-report measures first took place at diagnostic assessment prior to treatment, or within 2–4 weeks of commencing treatment at the latest. These measures were subsequently administered after 6 and 36 months.

2.4. Data analysis

Using the 10 key diagnostic variables, cluster analysis was conducted in a series of three steps using SLEIPNER (Bergman & El-Khouri, 1998). These steps have been detailed previously (Clinton et al., 2004) and resulted in a three-cluster classification of eating disorder patients (i.e., "generalised eating disorder", "anorexics," and "overeaters"). This three-cluster solution was used as a point of departure in the present study, allowing clusters to be compared with DSM-IV diagnoses on data pertaining to psychiatric comorbidity.

3. Results

3.1. Comparisons at intake

In order to aid comparisons of clusters and diagnoses, results of our previous cluster analysis for the Swedish sample are summarised in Table 1; the table presents data for the three clusters on the key diagnostic variables used for clustering. Comparisons between clusters and diagnoses on intake data were then made on the same key diagnostic variables, interview-based measures of co-morbidity, EDI-2 and the SCL-90 (63-item version). Results of one-way ANOVA (F-values, significance, and effect size) are presented in Table 2. Graphic comparisons of clusters and diagnoses are given for key diagnostic variables in Fig. 1a and b.

Table 1
Standard scores on key clinical variables in relation to three-cluster solution following non-hierarchical relocation analysis

Key diagnostic variables	Generalised eating disorder (GED) (N=216)	Clusters	
		Overeaters (N=193)	Anorexics (N=192)
BMI	.03	.65	−.68
Weight phobia	.39	−.36	−.07
Binge eating	.60	.28	−.96
Restriction	.47	−1.03	.51
Avoidance of fattening food	.48	−.88	.35
Vomiting	.68	.01	−.78
Laxative abuse	.32	−.25	−.11
Compulsive exercise	.26	−.49	.20
Amenorrhoea	−.24	−.33	.61
Disturbed body image	.64	−.63	.29

F-values were naturally significant across the board for virtually all of the key diagnostic variables used for cluster analysis; this was true for both clusters and diagnoses. What is most interesting, however, is a comparison of the effect sizes (h^2) of these variables in relation to clusters and diagnoses. A comparison of effect size gives an indication of how well the two approaches perform in their ability to account for variation in key eating disorder symptoms. When variables are examined in relation to DSM-IV diagnoses, high effect sizes were found in both samples for BMI, binge eating, vomiting, and amenorrhoea, which would suggest that the DSM system is, not surprisingly, primarily relying on these variables to explain the variance in diagnostic categories. When variables are examined in relation to clusters, effect sizes were, on the whole, even higher and more evenly distributed across a wider range of variables. The only instances where variables achieved notably higher effect sizes in relation to diagnoses compared to clusters were BMI. Slightly higher effect sizes in relation to diagnoses were found for binge eating and amenorrhoea. In contrast, effect sizes were considerably higher for restriction and avoidance of fattening food in relation to clusters as opposed to diagnoses in both samples. Higher effect sizes of a somewhat lesser magnitude were found for weight phobia, compulsive exercise, and disturbed body image.

On the interview-based assessment of co-morbidity, clusters were distinguished by markedly higher effect sizes in relation to suicidal behaviour and self-harm; when pair-wise differences were examined on this variable using Scheffé tests ($p<.05$), patients in the “generalised eating disorder” cluster scored significantly higher than “anorexics” and “overeaters.” On the EDI-2 and SCL-90 (63-item version), levels of significance and effect sizes were generally markedly greater for clusters compared to diagnoses. On the EDI, this was especially true of the psychological subscales. When pair-wise differences were examined for clusters using Scheffé tests, “overeaters” scored for the most part significantly lower than patients in the “generalised eating disorder” cluster, while the latter scored significantly higher than “overeaters” and “anorexics.” An important exception to this trend was found for the Bulimia subscale where “overeaters” scored significantly higher than “anorexics.” On the SCL-90 (63-item version), significant differences between diagnostic categories were only found for the depression subscale, while all between-cluster differences were significant. When pair-wise comparisons were made for clusters, patients in the “generalised eating disorder” cluster scored significantly higher than “overeaters” on all subscales. “Anorexics” scored

Table 2
Comparisons of clusters and DSM-IV diagnoses at initial assessment on one-way ANOVA

	Clusters[a]			DSM diagnoses[b]		
	F	*p*	h^2	*F*	*p*	h^2
Key diagnostic variables						
BMI	119.5	<.001	.29	153.8	<.001	.44
Weight phobia	32.3	<.001	.10	8.5	<.001	.04
Binge eating	247.8	<.001	.45	198.1	<.001	.50
Restriction	301.2	<.001	.50	33.8	<.001	.14
Avoidance of fattening food	175.8	<.001	.37	19.7	<.001	.09
Vomiting	168.0	<.001	.36	89.0	<.001	.31
Laxative abuse	19.2	<.001	.06	2.0	NS	.01
Compulsive exercise	38.5	<.001	.11	9.7	<.001	.05
Amenorrhoea	62.8	<.001	.17	65.0	<.001	.25
Disturbed body image	69.9	<.001	.19	11.4	<.001	.05
Interview measures of co-morbidity						
Drug and alcohol problems	6.3	<.01	.02	3.9	<.01	.02
Suicidality	26.0	<.001	.08	2.5	NS	.01
Relationship problems	1.8	NS	.01	2.2	NS	.01
EDI-2						
Drive for thinness	32.5	<.001	.10	15.2	<.001	.07
Bulimia	165.9	<.001	.36	76.6	<.001	.28
Body dissatisfaction	15.7	<.001	.05	13.3	<.001	.06
Ineffectiveness	17.0	<.001	.05	4.6	<.01	.02
Perfectionism	9.2	<.001	.03	0.7	NS	.00
Interpersonal distrust	7.7	<.001	.02	0.9	NS	.00
Introceptive awareness	18.2	<.001	.06	2.7	<.05	.01
Maturity fears	8.4	<.001	.03	6.5	<.001	.03
Asceticism	14.9	<.001	.05	3.3	<.05	.02
Disturbed impulse regulation	7.9	<.001	.03	2.8	<.05	.01
Social insecurity	7.3	<.001	.02	2.5	NS	.01
Total ED subscales	64.4	<.001	.18	39.1	<.001	.17
Total psychological subscales	22.5	<.001	.07	3.9	<.01	.02
Total score	40.0	<.001	.12	10.4	<.001	.05
SCL 63						
Somaticism	20.0	<.001	.06	1.5	NS	.01
Obsession–compulsion	8.3	<.001	.03	1.7	NS	.01
Interpersonal sensitivity	19.0	<.001	.07	1.2	NS	.01
Depression	12.4	<.001	.04	1.9	NS	.01
Anxiety	14.3	<.001	.05	3.7	<.05	.02
Anger	5.7	<.01	.02	0.5	NS	.00
Symptom index	20.5	<.001	.06	2.3	NS	.01

[a] Clusters: generalised eating disorder (GED), overeaters, anorexics.
[b] DSM diagnoses: AN, BN, BED, EDNOS.

significantly higher than “overeaters” on Somaticism, and Interpersonal sensitivity. “Anorexics” also scored significantly lower than “generalised eating disorder” patients on Somaticism, Interpersonal sensitivity, Depression, Anxiety, and Anger. Graphic comparisons between clusters and diagnoses on the EDI-2 are presented in Fig. 2a and b, while comparisons on the SCL-90 (63-item version) are presented in Fig. 3a and b.

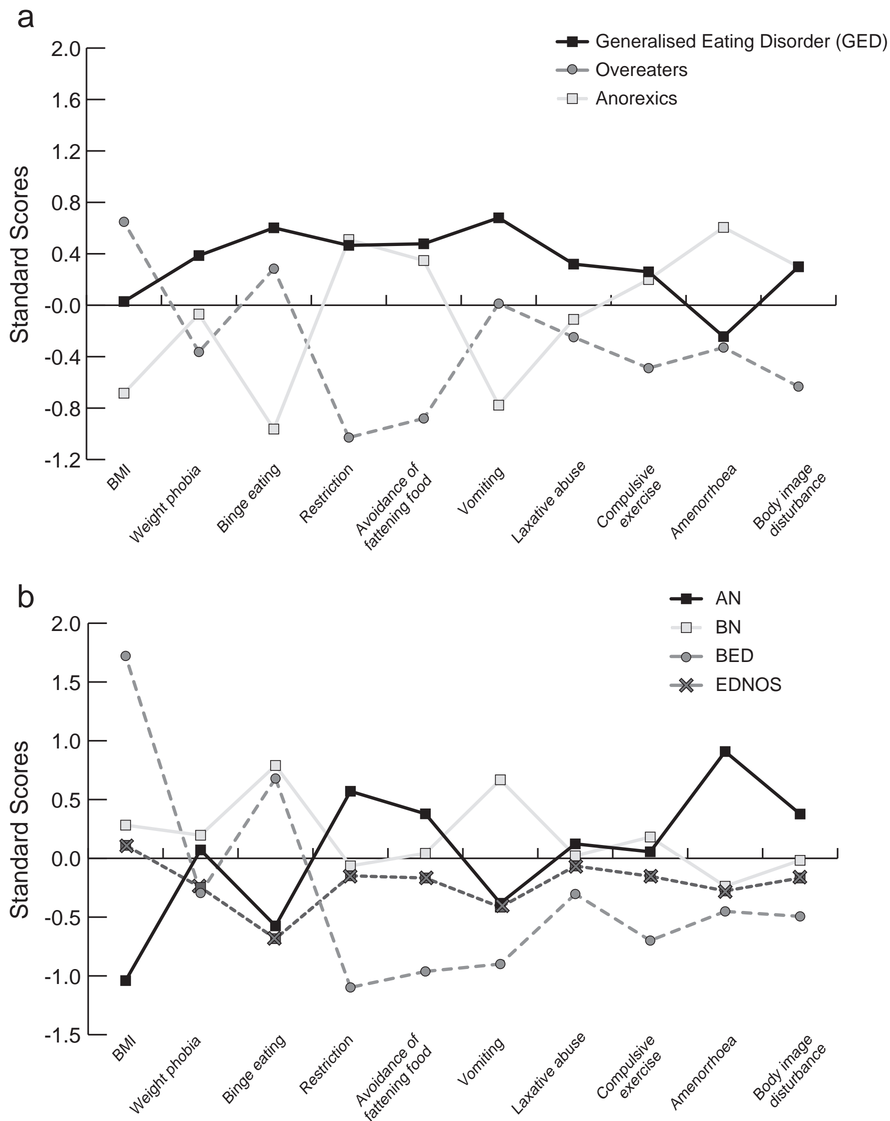

Fig. 1. (a) Comparisons of clusters on key diagnostic variables. (b) Comparisons of DSM-IV diagnoses on key diagnostic variables.

3.2. Comparisons after 6 and 36 months

Further comparisons of clusters and diagnoses were made by examining patients who had been followed up after 6 and 36 months. Results of one-way ANOVA (F-values, significance, and effect size) are presented in Tables 3 (after 6 months) and 4) (after 36 months).

After 6 months, levels of significance and effect sizes were diminished for both clusters and diagnoses but were still considerably greater for clusters compared to diagnoses. Between-cluster differences were still significant on all key diagnostic variables, but only half of the between-diagnosis differences. Effect

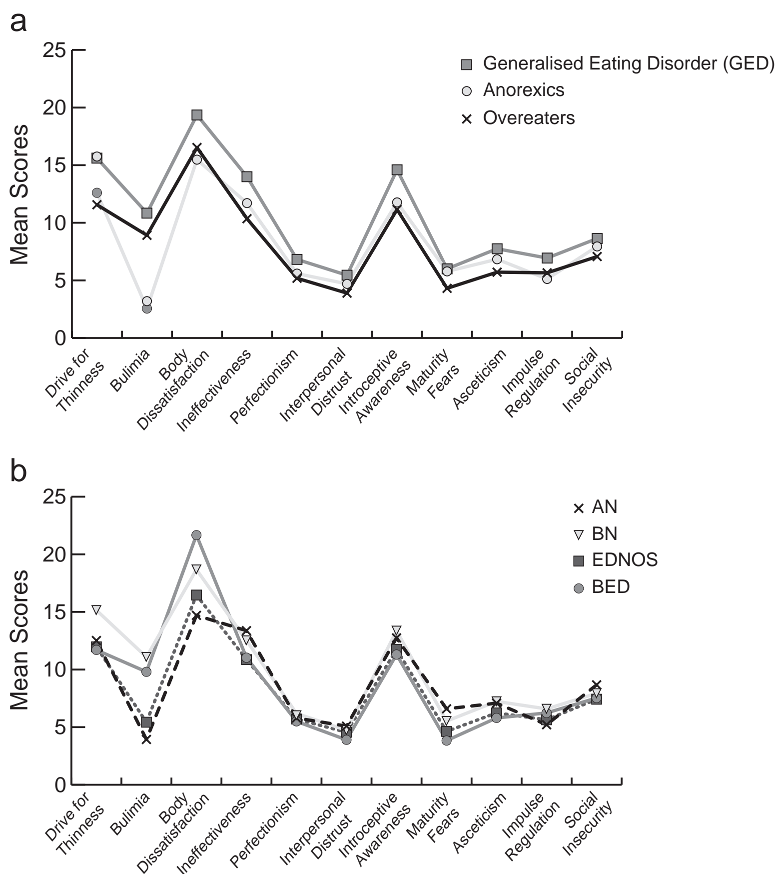

Fig. 2. (a) Mean scores for clusters on the EDI-2 at initial assessment. (b) Mean scores for diagnoses on the EDI-2 at initial assessment.

sizes were greater for diagnoses on BMI and amenorrhoea, but roughly at least double for clusters on the remaining key variables. On the SCL-90 (63-item version), one-way ANOVA was significant on all subscales for clusters but none of the comparisons attained significance for diagnoses. After 36 months, the patterns that had been marked at assessment and still present after 6 months had been largely diminished. On key diagnostic variables, effect sizes were still greater for diagnoses on BMI and amenorrhoea, but still roughly twice as large for clusters on the remaining variables. On the EDI-2, significant between-cluster differences were still present on the Bulimia subscale, while significant between-diagnosis differences were found on Maturity fears. On the SCL-90 (63-item version), Anger was still significant for clusters.

4. Discussion

We have attempted to clarify the nosology of eating disorders by comparing the clinical utility of DSM-IV diagnoses with an alternative classification derived statistically in our previous work (Clinton et al.,

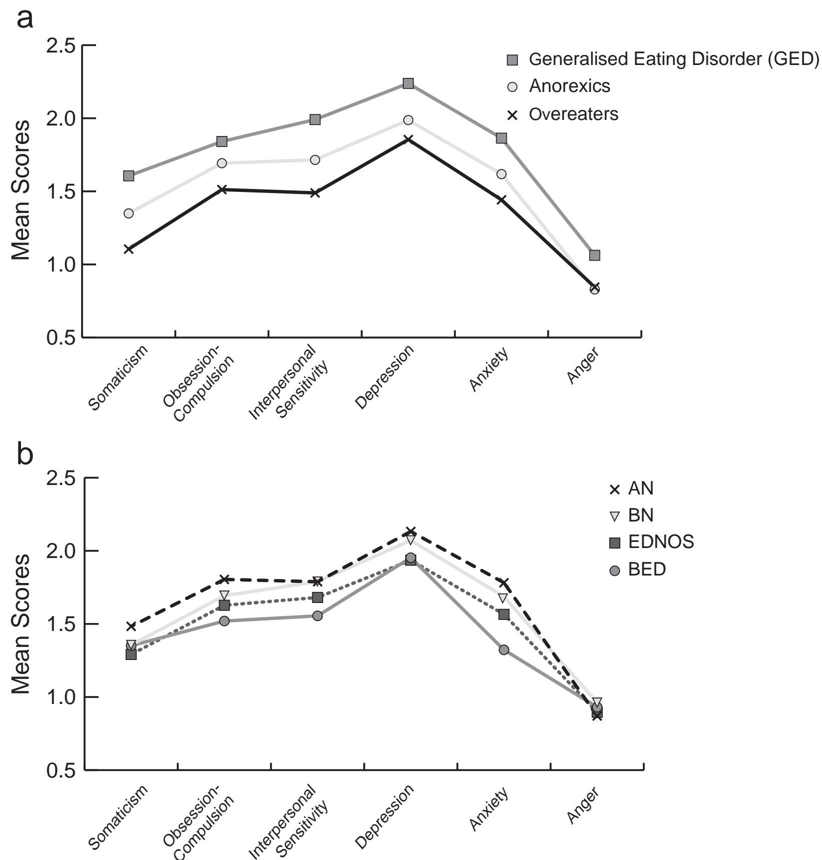

Fig. 3. (a) Mean scores for clusters on the SCL-90(63-item version) at initial assessment. (b) Mean scores for diagnoses on the SCL-90(63-item version) at initial assessment.

2004). In our previous study, using a large heterogeneous sample of eating disorder patients with a broad spectrum of diagnoses from both Sweden and the UK, we found clinically recognisable clusters of patients. These were labelled "anorexics," "overeaters," and "generalised eating disorder." Compared to DSM-IV diagnoses, clusters placed greater emphasis on psychological and behavioural features of eating disorders and generated a large category of patients who were suffering from a cross section of eating disorder symptoms (i.e., those classified as "generalised eating disorder"). In particular, patients who were classified as "generalised eating disorder" tended to present with a more distinct array of symptoms compared to patients diagnosed as EDNOS. The former constituted a group with high levels of symptomatology in almost every respect, except for weight and menstrual functioning, while the latter tended to constitute a group of patients without distinguishing features on any of the key diagnostic variables.

The present study takes our previous work a step further and suggests that cluster analysis did not just generate a classification of considerable heuristic value. It also generated categories that demonstrated a higher degree of utility than conventional diagnoses. The higher degree of association between key eating disorder symptoms and clusters as opposed to diagnoses suggests that clusters "out-performed" diagnoses in their ability to account for variability in important symptoms. This was the case at intake, as

Table 3
Comparisons of clusters and DSM-IV diagnoses at 6-month follow-up on one-way ANOVA

	Clusters[a]			DSM diagnoses[b]		
	F	*p*	h^2	*F*	*p*	h^2
Key diagnostic variables						
BMI	44.1	<.001	.25	52.5	<.001	.37
Weight phobia	5.0	<.01	.03	1.4	NS	.01
Binge eating	34.1	<.001	.17	12.3	<.001	.10
Restriction	8.8	<.001	.05	2.6	NS	.02
Avoidance of fattening food	16.9	<.001	.09	1.5	NS	.01
Vomiting	32.3	<.001	.16	10.5	<.001	.08
Laxative abuse	4.0	<.05	.02	0.9	NS	.01
Compulsive exercise	3.3	<.05	.02	1.4	NS	.01
Amenorrhoea	33.0	<.001	.17	43.1	<.001	.28
Disturbed body image	6.4	<.01	.04	2.6	<.05	.02
EDI-2						
Drive for thinness	11.5	<.001	.06	3.6	<.05	.03
Bulimia	20.5	<.001	.11	5.9	<.001	.05
Body dissatisfaction	4.1	<.05	.02	3.5	<.05	.03
Ineffectiveness	4.9	<.01	.03	3.7	<.05	.03
Perfectionism	2.0	NS	.01	0.4	NS	.00
Interpersonal distrust	5.2	<.01	.03	1.8	NS	.01
Introceptive awareness	12.2	<.001	.06	2.9	<.05	.02
Maturity fears	5.0	<.01	.03	4.8	<.01	.04
Asceticism	4.7	<.01	.03	1.4	NS	.01
Disturbed impulse regulation	2.6	NS	.01	0.9	NS	.01
Social insecurity	2.4	NS	.01	2.0	NS	.02
Total ED subscales	12.0	<.001	.07	4.8	<.01	.04
Total psychological subscales	7.6	<.001	.04	3.6	<.05	.03
Total score	9.8	<.001	.05	3.5	<.05	.03
SCL-63						
Somaticism	5.5	<.01	.03	1.6	NS	.01
Obsession–compulsion	5.7	<.01	.03	1.1	NS	.01
Interpersonal sensitivity	7.1	<.001	.04	2.3	NS	.02
Depression	7.1	<.001	.04	2.5	NS	.02
Anxiety	9.3	<.001	.05	1.8	NS	.01
Anger	6.1	<.01	.02	1.6	NS	.01
Symptom index	8.9	<.001	.06	1.5	NS	.01

[a] Clusters: generalised eating disorder (GED), overeaters, anorexics.
[b] DSM diagnoses: AN, BN, BED, EDNOS.

well as at short-term follow-up after 6 months. It was also the case in relation to both a wider range of eating disorder psychopathology and important aspects of comorbidity. Even when variables that were not used for cluster analysis were compared, a more distinct and homogeneous pattern of between-group differences emerged, and effect sizes were greater in relation to clusters as opposed to diagnoses.

The ability of the cluster approach to generate a more homogeneous and powerful classification may be due to two important factors. Firstly, cluster analysis attached greater importance to a wider range of features, in particular psychological and behavioural characteristics of eating disorders in order to classify cases, as opposed to a more singular emphasis of weight, menstrual functioning, and binge eating.

Table 4
Comparisons of clusters and DSM-IV diagnoses at 36-month follow-up on one-way ANOVA

	Clusters[a]			DSM diagnoses[b]		
	F	*p*	h^2	*F*	*p*	h^2
Key diagnostic variables						
BMI	119.5	<.001	.29	153.8	<.001	.44
Weight phobia	32.3	<.001	.10	8.5	<.001	.04
Binge eating	247.8	<.001	.45	198.1	<.001	.50
Restriction	301.2	<.001	.50	33.8	<.001	.14
Avoidance of fattening food	175.8	<.001	.37	19.7	<.001	.09
Vomiting	168.0	<.001	.36	89.0	<.001	.31
Laxative abuse	19.2	<.001	.06	2.0	NS	.01
Compulsive exercise	38.5	<.001	.11	9.7	<.001	.05
Amenorrhoea	62.8	<.001	.17	65.0	<.001	.25
Disturbed body image	69.9	<.001	.19	11.4	<.001	.05
Interview measures of co-morbidity						
Drug and alcohol problems	6.3	<.01	.02	3.9	<.01	.02
Suicidality	26.0	<.001	.08	2.5	NS	.01
Relationship problems	1.8	NS	.01	2.2	NS	.01
EDI-2						
Drive for thinness	2.6	NS	.02	0.9	NS	.01
Bulimia	4.2	<.05	.03	0.1	NS	.00
Body dissatisfaction	0.8	NS	.00	0.5	NS	.00
Ineffectiveness	0.2	NS	.00	2.0	NS	.02
Perfectionism	2.1	NS	.01	1.1	NS	.01
Interpersonal distrust	0.8	NS	.00	0.9	NS	.01
Introceptive awareness	1.9	NS	.01	0.3	NS	.00
Maturity fears	0.9	NS	.01	4.5	<.01	.04
Asceticism	0.4	NS	.00	0.9	NS	.01
Disturbed impulse regulation	0.3	NS	.00	0.4	NS	.00
Social insecurity	0.7	NS	.00	0.8	NS	.01
Total ED subscales	1.7	NS	.01	0.2	NS	.00
Total psychological subscales	0.3	NS	.00	1.1	NS	.01
Total score	0.1	NS	.00	0.3	NS	.00
SCL-63						
Somaticism	2.6	NS	.02	0.6	NS	.00
Obsession–compulsion	2.7	NS	.02	1.4	NS	.01
Interpersonal sensitivity	1.5	NS	.01	1.5	NS	.01
Depression	0.3	NS	.00	0.7	NS	.01
Anxiety	1.7	NS	.01	0.7	NS	.01
Anger	3.1	<.05	.02	0.1	NS	.00
Symptom index	1.7	NS	.01	0.7	NS	.01

[a] Clusters: generalised eating disorder (GED), overeaters, anorexics.
[b] DSM diagnoses: AN, BN, BED, EDNOS.

Moreover, cluster analysis put comparatively greater emphasis on restriction of eating, avoidance of fattening foods, weight phobia, disturbed body image, compulsive exercise and laxative abuse. Secondly, cluster analysis succeeded in reallocating patients consigned to the residual diagnostic category of EDNOS to more relevant categories. This in itself is of importance since this sizable group of patients has proved largely unamenable to classification by conventional means. On the whole, our results raise the question of

making changes to prevailing diagnostic systems. Essentially, we are faced with two alternatives to the status quo. On the one hand, we might consider revisions and refinements of prevailing criteria in the direction of less specificity. On the other, we might be tempted by a more radical approach and contemplate moving towards a new system that is more phenomenologically and clinically relevant.

Over the years, the DSM system has tended to revise the criteria of eating disorders in terms of increased precision and specificity of AN and BN. Although this has undoubtedly helped to distinguish AN from BN, it may have also resulted in definitions of these syndromes that are too narrow and not therefore in touch with clinical reality. Our results suggest that greater clinical utility might in fact be achieved by loosening some of the criteria used for diagnosing AN and BN. For example, in AN, we might consider discarding Criterion D (i.e., the absence of at least three consecutive menstrual cycles in females). Alternatively, instead of requiring current amenorrhoea for a diagnosis of AN, this criterion might be added to a list of other key criteria of which a specified number of these would be required to establish a diagnosis. Criterion B (i.e., intense fear of gaining weight or becoming fat, even though underweight) could be amended to reflect a common criterion for all eating disorders, such as an "over-investment of eating restraint." In BN, we might consider being less restrictive in relation to questions of frequency and duration. Criterion C (i.e., occurrence of binge eating and compensatory behaviours at least twice a week for 3 months) might be dropped or amended to specify a minimal frequency without a minimal duration. Criterion D (i.e., the undue influence of body shape and weight for self-evaluation) could also be changed to reflect a common criterion for all eating disorders.

Recent empirical attempts at revisions and refinements of prevailing diagnostic criteria have focused on widening criteria. Weston and Harnden-Fischer (2001) have advocated going beyond axis 1 to include relevant personality dimensions to better classify eating disorders. Studies by Bulik et al. (2000) and Crow, Agras, Halmi, Mitchell, and Kraemer (2002), suggest that present diagnostic classes are too narrowly defined, and that the relocation of a proportion of the patients diagnosed as EDNOS might be aided by the widening of criteria for the two primary syndromes of AN and BN. Andersen, Bowers, and Watson (2001) experimented with the specific revisions of criteria. They eliminated the criterion for amenorrhoea in AN and required a 20% loss of weight rather than a weight level 15% under normal weight. For BN, they eliminated the criteria concerning the frequency and duration of binge eating and compensatory behaviour. These changes resulted in a reduction of the EDNOS group to less than 20% of its original size, as well as an increase in the AN group by more than 50%, while the BN group remained roughly the same size. In a replication of the study by Andersen et al. using data from our own CO-RED project ($N=806$), the EDNOS group was reduced to approximately a third of its original size, while the AN group increased by 150%, the BN group was reduced by 15%, and the BED group was reduced by more than half (Norring, 2002). In terms of symptoms, the most notable result of this reallocation was that the AN group came to comprise a number of patients who were not obviously underweight, and the EDNOS patients who remained in the EDNOS category reported markedly lower scores on the EDI compared to patients with other eating disorder diagnoses. Changes such as these in specific criteria for eating disorders might help in some ways, but it is doubtful whether they would be met with enthusiasm by clinicians, or whether they would result in a more clinically relevant system of diagnosis.

We tend to agree with Beumont, Garner, and Touyz (1994) when they reviewed proposed changes for DSM-IV and concluded, "The solution of the diagnostic muddle is unlikely to come from tinkering with the criteria once again. The composite package approach, listing and defining specific groups of symptoms, has exhausted its usefulness." This brings us to the second alternative to the status quo involving a new approach to the diagnosis of eating disorders. In the present study, the largest group of patients fell into the

"generalised eating disorder" cluster. The predominance of this cluster could be taken as suggesting the need to amend diagnostic systems to accommodate such a category. Although such a category is arguably both clinically useful and phenomenologically accurate with its close attention to core features of eating disorders, the "generalised eating disorder" cluster can also be interpreted differently. The cluster may in fact be an apt reminder that the time is ripe to seriously contemplate a move away from increased sub-categorising of eating disorders toward defining a diagnosis of eating disorder per se. Such a move, it should be pointed out, would not necessarily involve abandoning the sub-type approach, but complementing it with categories that emphasise the common characteristics of eating disorders.

Waller (2005) argues that we need to "abandon our efforts to create subcategories and get used to working with the core features of eating disorders." Attempts are being made in this direction. Recently, Fairburn and Walsh (2002) proposed 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." Whether the emphasis should be placed on weight concern or behaviour intended to control weight is, however, open to debate. Palmer (1993) argued that although weight concern has clear applicability to the diagnosis of AN and BN, it has important problems as a general criterion. He drew attention to what he called "eating restraint that is over-invested." Amongst other things, he argued that eating restriction (or attempts at eating restriction) appears to be a universal finding in the histories of patients presenting for treatment of an eating disorder. Once we have defined eating disorders, it will be possible to specify typical clusters of symptoms, or even dimensions of symptoms, and thereby also identify meaningful subgroups. Clinically, we are already seeing the consequences of an increased focus on the core features of eating disorders. Fairburn, Cooper, and Shafran (2003) have put forward a "transdiagnostic theory" of eating disorders, which is based around the idea that eating disorders, regardless of subtype, are maintained by the same psychopathological processes. He argues that this has important clinical implications and has developed a modified and flexible form of CBT that can be applied across specific diagnoses.

Solving the problem of the classification of eating disorders will require both further research and careful deliberation. At present, there is increasing evidence that the way ahead will require a move away from increased subtyping and toward a definition of eating disorder per se. It will be important for such a definition to pay close attention to the psychological features of these disorders. Empirically, there is still important work to be done. Future studies need to examine the issue of utility in relation to other variables (e.g., personality and outcome measures), as well as explore other variables than those used for clustering in the present study. Paying increased attention to the interaction of personality disorders and eating disorder symptoms in relation to classification will be a particularly important area of future research. When we have gathered data on the merits of alternative schemes of classification, it will be possible to define and test these alternatives and compare them to conventional criteria. Together, these steps will help move us toward a rational system of classifying eating disorders based on more accurate phenomenology and greater clinical utility.

References

Andersen, A. E., Bowers, W. A., & Watson, T. (2001). A slimming program for eating disorders not otherwise specified: Reconceptualizing a confusing, residual diagnostic category. *The Psychiatric Clinics of North America*, *24*, 271–280.

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

Bergman, L., & El-Khouri, B. M. (1998). *SLEIPNER: A statistical package for pattern-oriented analysis*. Stockholm University: Department of Psychology.

Beumont, P. J. V., Garner, D. M., & Touyz, S. W. (1994). Diagnoses of eating or dieting disorders: What may we learn from past mistakes? *International Journal of Eating Disorders*, *16*, 349–362.

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.

Clinton, D. N., 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. N., & Norring, C. (1999). The Rating of Anorexia and Bulimia (RAB) interview: Development and preliminary validation. *European Eating Disorders Review*, *7*, 362–371.

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.

Derogatis, L. R., Lipman, R. S., & Covi, L. (1973). SCL-90: An outpatient psychiatric rating scale-preliminary report. *Psychopharmacology*, *9*, 13–28.

Everitt, B. S., & Landau, S. (1998). The use of multivariate statistical methods in psychiatry. *Statistical Methods in Medical Research*, *7*, 253–277.

Everitt, B. S., Landau, S., & Leese, M. (2001). *Cluster analysis* (4th ed.). London: Arnold.

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., & 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 K. D. Brownell, & C. G. Fairburn (Eds.), *Eating disorders and obesity: A comprehensive handbook*, (2nd ed.). New York: Guilford Press.

Garner, D. M., Olmsted, M. P., & Polivy, J. (1983). Development and validation of a mulidimensional Eating Disorders Inventory for anorexia nervosa and bulimia. *International Journal of Eating Disorders*, *2*, 15–34.

Grilo, C. M., Masheb, R. M., & Berman, R. M. (2001). Subtyping women with bulimia nervosa along dietary and negative affect dimensions: A replication in a treatment-seeking sample. *Eating and Weight Disorders*, *6*, 53–58.

Hay, P. J., Fairburn, C. G., & Doll, H. A. (1996). The classification of bulimic eating disorders: A community-based cluster analysis. *Psychological Medicine*, *26*, 801–812.

Keel, P., Fichter, M., Quadflieg, N., Bulik, C., Baxter, M., Thornton, L., et al. (2004). Application of a latent class analysis to empirically define eating disorder phenotypes. *Archives of General Psychiatry*, *61*, 192–200.

Mizes, J. S., & Sloan, D. M. (1998). An empirical analysis of eating disorder, not otherwise specified: Preliminary support for a distinct sub-group. *International Journal of Eating Disorders*, *23*, 233–242.

Nevonen, L., Broberg, A., Clinton, D., & Norring, C. (2003). A measure for the assessment of eating disorders: Reliability and validity studies of the Rating of Anorexia and Bulimia interview—revised version (RAB-R). *Scandinavian Journal of Psychology*, *44*, 303–310.

Norring, C. (2002). Can we contribute to the development of clinically more relevant and meaningful eating disorder diagnoses? *Paper presented at the 4th Nordic Congress on Eating Disorders, Lund, Sweden, May 28–31*.

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.

Palmer, R. L. (2003). Concepts of eating disorders. In U. Schmidt, J. Treasure, & E. van Furth (Eds.), *Handbook of eating disorders*, (2nd ed.). Chichester: John Wiley and Sons.

Russell, G. F. M. (1979). Bulimia nervosa: An ominous variant of anorexia nervosa? *Psychological Medicine*, *9*, 429–448.

Sloan, D. M., Mizes, J. S., & Epstein, E. M. (2005). Empirical classification of eating disorders. *Eating Behaviours*, *6*, 53–62.

Stice, E., & Agras, W. S. (1999). Subtyping bulimic women along dietary restraint and negative affect dimensions. *Journal of Consulting and Clinical Psychology*, *67*, 460–469.

Turner, H., & Bryant-Waugh, R. (2004). Eating disorder not otherwise specified (EDNOS): Profiles of patients presenting at a community eating disorder service. *European Eating Disorders Review*, *12*, 18–26.

Van der Ham, T., Meulman, J. J., van Strien, D. C., & van Engeland, H. (1997). Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *British Journal of Psychiatry*, *170*, 363–368.

Waller, G. (2005). Psychological perspectives on atypical diagnoses in the eating disorders. In C. Norring, & R. L. Palmer (Eds.), *EDNOS: Eating disorders not otherwise specified—scientific and clinical perspectives on the other eating disorders*. London: Brunner-Routledge.

Weston, D., & Harnden-Fischer, J. (2001). Personality profiles in eating disorders: Rethinking the distinction between axis I and axis II. *American Journal of Psychiatry*, *158*, 547–562.

WHO (1992). *International classification of diseases* (10th ed.). Geneva: World Health Organization.