Psychological Medicine (2010), **40**, 1203–1213. © Cambridge University Press 2009
doi:10.1017/S003329170999122X

ORIGINAL ARTICLE

A twin study of specific bulimia nervosa symptoms

S. E. Mazzeo[1,2,3]*, K. S. Mitchell[1,3], C. M. Bulik[4], S. H. Aggen[3,5], K. S. Kendler[3,5] and M. C. Neale[3,5,6]

[1] *Department of Psychology, Virginia Commonwealth University, Richmond, VA, USA*
[2] *Department of Pediatrics, Virginia Commonwealth University, Richmond, VA, USA*
[3] *Virginia Institute for Psychiatric & Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA*
[4] *Departments of Psychiatry and Nutrition, University of North Carolina, Chapel Hill, NC, USA*
[5] *Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, USA*
[6] *Department of Human Genetics, Virginia Commonwealth University, Richmond, VA, USA*

Background. Twin studies have suggested that additive genetic factors significantly contribute to liability to bulimia nervosa (BN). However, the diagnostic criteria for BN remain controversial. In this study, an item-factor model was used to examine the BN diagnostic criteria and the genetic and environmental contributions to BN in a population-based twin sample. The validity of the equal environment assumption (EEA) for BN was also tested.

Method. Participants were 1024 female twins (MZ $n=614$, DZ $n=410$) from the population-based Mid-Atlantic Twin Registry. BN was assessed using symptom-level (self-report) items consistent with DSM-IV and ICD-10 diagnostic criteria. Items assessing BN were included in an item-factor model. The EEA was measured by items assessing similarity of childhood and adolescent environment, which have demonstrated construct validity. Scores on the EEA factor were used to specify the degree to which twins shared environmental experiences in this model.

Results. The EEA was not violated for BN. Modeling results indicated that the majority of the variance in BN was due to additive genetic factors. There was substantial variability in additive genetic and environmental contributions to specific BN symptoms. Most notably, vomiting was very strongly influenced by additive genetic factors, while other symptoms were much less heritable, including the influence of weight on self-evaluation. These results highlight the importance of assessing eating disorders at the symptom level.

Conclusions. Refinement of eating disorder phenotypes could ultimately lead to improvements in treatment and targeted prevention, by clarifying sources of variation for specific components of symptomatology.

Received 6 May 2008; Revised 28 February 2009; Accepted 3 March 2009; First published online 12 October 2009

Key words: Bulimia nervosa, heritability, item-factor model, twins.

Introduction

Bulimia nervosa (BN) is an eating disorder with debilitating physical and psychological effects (Becker *et al*. 1999). Moreover, BN severely affects the quality of life of both affected individuals (Simon *et al*. 2005) and their families (Winn *et al*. 2007). Winn *et al*. (2007) found that the majority of caregivers of individuals with BN experienced significant psychological distress of their own; caregivers' distress was comparable to that of caregivers of individuals with psychosis. The economic costs of BN are also astounding; for example, Simon and colleagues found that costs for BN in Germany were €10 million in 1998 alone. Further, these expenditures are likely underestimated, as only a small proportion of affected individuals seek treatment (Cachelin *et al*. 2000; Simon *et al*. 2005). Thus, BN has significant costs for individuals, families and society.

Despite the devastating effects of BN, and its relatively high prevalence among women in Western society (approximately 1.5% among women in the United States; Hudson *et al*. 2007), many questions remain about the etiology of this disorder. Emphasis has long been placed on the influence of familial interaction styles and socio-cultural factors on the etiology of eating disorders (for reviews, see le Grange, 2005; Becker & Fay, 2006). However, in the last two decades, twin studies have suggested that genetic factors significantly influence BN symptomatology (Kendler *et al*. 1991, 1995; Bulik *et al*. 1998, 2000; Sullivan *et al*. 1998; Bulik, 2005).

Nevertheless, the validity of the conclusions drawn from twin research, like all psychological research, is dependent upon the quality of both the operationalization of the constructs (or latent variables) of interest and the measurement approach used to assess these

* Address for correspondence: Dr S. E. Mazzeo, Department of Psychology, Virginia Commonwealth University, PO Box 842018, Richmond, VA 23284–2018, USA.
(Email: semazzeo@vcu.edu)

constructs. The operationalization of BN can be problematic, particularly because diagnostic criteria have changed across revisions of the DSM and, currently, the BN criteria, like those of all eating disorders, are hotly debated (e.g. Bulik *et al*. 2007; Wilfley *et al*. 2007). Previous studies of the heritability of BN (Kendler *et al*. 1991, 1995; Bulik *et al*. 1998; Sullivan *et al*. 1998) have generally investigated this disorder using a 'sum score' approach, in which items assessing DSM criteria are aggregated to form a diagnostic composite. However, more recent research has suggested that, because individual items are rarely pure indicators of a single latent factor, sum scores might confound measurement of the latent trait by ignoring item-specific variance components (Neale *et al*. 2005). For example, if the amount of heritability differed greatly for individual items, the heritability of their sum score could be either over- or underestimated. Further, individual symptoms of a disorder such as BN are likely differentially influenced by genetic and environmental factors (Bulik *et al*. 2007). Thus, the current study used an item-factor approach, including all items assessing DSM-IV criteria for BN, to assess additive genetic, common environmental and unique environmental influences on this disorder.

In addition to the construct validity of BN measures, the validity of the twin model itself is dependent upon non-violation of the equal environment assumption (EEA). The EEA posits that MZ and DZ twins share to an equal extent those environmental influences that are: (*a*) etiologically relevant to the trait or diagnosis under investigation; (*b*) not elicited by the twins (Kendler *et al*. 1993; Klump *et al*. 2000*b*). The validity of the EEA is trait-specific and must be examined for each diagnosis of interest. If the EEA is violated, then greater MZ (*versus* DZ) correlations could be a result of non-genetic effects and heritability may be over-estimated (Fairburn *et al*. 1999). Previous studies have examined the validity of the EEA in eating disorders (e.g. Kendler *et al*. 1993; Hettema *et al*. 1995; Bulik *et al*. 1998; Klump *et al*. 2000*a*), and results have generally suggested that this assumption was not violated, with a few exceptions (e.g. Hettema *et al*. 1995). However, no previous studies have used a quantitative factor-score measure of twins' environments, such as that recently evaluated by Mitchell *et al*. (2007), as a definition variable, or moderator, within a biometric model. This approach enables a direct examination of the moderating influence of rearing twins similarly (i.e. EEA factor scores) during childhood and adolescence. This separate assessment of childhood and adolescent environment seems particularly relevant to eating disorders, as research has suggested that the influence of genetic and environmental factors changes over the course of development, particularly following puberty in girls (e.g. Klump *et al*. 2000*b*).

Thus, the present study had two aims. First, we examined genetic and environmental contributions to variance specific to each BN symptom (as well as the overall BN liability) using an item-factor modeling approach. This method resolves many problems associated with summing items assessing symptoms of an overall diagnosis. Use of this modeling method enables investigation of: (*a*) how strongly BN symptoms (observed variables) are related to the latent trait (BN); and (*b*) the proportions of variance due to genetic and environmental influences for the latent trait and individual items (Neale *et al*. 2005). Second, we investigated the validity of the EEA, as applied to BN, in a population-based sample of female twins. The following sections briefly review: (1) evidence for genetic influences on BN symptoms; (2) previous investigations of the validity of the EEA for BN.

Heritability of bulimia nervosa symptoms

Binge eating

Previous studies have found that additive genetic effects account for significant variance in binge eating, with heritability estimates ranging from 41% to 82% (Bulik *et al*. 1998, 2007; Sullivan *et al*. 1998; Reichborn-Kjennerud *et al*. 2003.). However, operationalization of this construct has been critiqued due to difficulties interpreting specific aspects of the definition, including 'loss of control', 'large amount' and 'short period of time' (Bulik *et al*. 2007; Wilfley *et al*. 2007). Thus, estimates of genetic and environmental influences might be confounded by these limitations. Use of item-factor modeling can help address this interpretive issue by providing separate estimates for items assessing binge eating and perceived loss of control.

Compensatory behaviors

A range of compensatory behaviors could be included under this criterion (including vomiting, laxative/diuretic abuse and excessive exercise); however, the most frequently studied is self-induced vomiting. This symptom is significantly influenced by additive genetic factors [e.g. $a^2 = 0.72$, 95% confidence interval (CI) 0.55–0.88; Sullivan *et al*. 1998]. Moreover, as Bulik *et al*. (2007) noted in their review, self-induced vomiting appears to be more reliably measured than binge eating, which can be challenging to assess due to the difficulties involved in interpreting this criterion discussed in the previous paragraph.

https://doi.org/10.1017/S003329170999122X Published online by Cambridge University Press

Binge eating frequency and duration

Twin studies have not examined heritability of the BN criterion, requiring that binge eating occurs for at least 3 months (Bulik *et al*. 2007). However, researchers have found few, if any, differences between participants who binged once per week and those who binged more frequently. For example, Spoor *et al*. (2007) found no differences in degree of psychosocial impairment or mental health care utilization between women who engaged in binge eating and compensatory behavior at subthreshold (one to seven times per month) and threshold (eight or more times per month) levels. These authors also found that duration of BN symptomatology was not associated with either psychosocial impairment or mental health care utilization. Based on these results, Spoor *et al*. conclude that the DSM-IV frequency criterion for BN may be excessive.

Undue influence of weight and shape on self-evaluation

As Mazzeo *et al*. (2009) have noted, findings regarding the influence of genetic and environmental factors on weight concern or the undue influence of appearance on self-evaluation have been mixed, with some studies indicating that this construct is significantly influenced by shared and unshared environment (e.g. Wade *et al*. 1998; Reichborn-Kjennerud *et al*. 2004), and others finding that additive genetic factors contribute significantly (Keski-Rahkonen *et al*. 2005). These inconsistencies may be related to variability in the measures used and subtle differences in the latent constructs assessed (Bulik *et al*. 2007).

In the only study to date to use an item-factor modeling approach to assess the genetic and environmental contributions of variance to the influence of weight on self-evaluation, Mazzeo *et al*. (2009) found evidence of moderate influences of additive genetic factors ($a^2 = 0.29$) on liability to this construct within a sample of Norwegian female twins. Unshared environment was the strongest contributor to variance in this construct ($e^2 = 0.62$), while the influence of shared environment was small ($c^2 = 0.08$). However, in addition to the measurement and construct validity issues noted above, results were also likely influenced by the questionnaire format. Specifically, because this was a population-based survey of numerous disorders, gateway items were used to reduce participant burden. Thus, only a subgroup of the sample, namely, those with a low enough body mass index (BMI) to be considered for the anorexia nervosa (AN) diagnosis, were included in the analyses. The authors concluded that their results suggest that genetic and environmental factors might operate differently within individuals already at a low BMI, compared with the general population. Moreover, these findings highlight the importance of studying specific symptoms within subgroups of interest, as heritability may differ across subgroups that vary in their risk of the disorder under investigation.

The equal environment assumption in twin studies of bulimia nervosa

The second aim of the present study was to investigate further the validity of the EEA, applied to BN, in a population-based sample of female twins. A variety of twin environment measures have been used in previous studies. For example, Kendler *et al*. (1993) conducted one of the first tests of the EEA in twin studies of BN and found that parents' perceptions of their twins' zygosity had no effect on rates of broadly defined BN in a population-based sample.

Similar results (based on the same sample) were found by Sullivan *et al*. (1998), who measured six aspects of common environment: childhood treatment (e.g. being dressed alike); co-socialization during childhood and adolescence; similitude (e.g. emphasis placed by parents and teachers on twins' similarity); physical similarity (rated using photographs); degree of adult contact; parents' ratings of the degree to which they emphasized twins' similarities. No EEA violations were found for either binge eating or vomiting.

However, other investigations have raised concerns about the validity of the EEA in twin studies of eating disorders. Bulik *et al*. (1998) investigated the EEA (using the measures described above) in a study of binge eating and broadly defined BN. They did not find any EEA violations for binge eating; however, twin co-socialization was associated with BN concordance. Hettema *et al*. (1995) examined the validity of the EEA in a range of psychiatric disorders, including broadly defined BN. EEA was evaluated using ratings of twins' physical similarity in photographs, as well as parents' and twins' assessments of physical similarity. Results indicated that twins who looked more alike were more vulnerable to BN.

Nonetheless, limitations of Hettema *et al*.'s study (1995) have been noted (Klump *et al*. 2000*a*). Klump and colleagues note that physical similarity and BN status might have been confounded in the Hettema *et al*. study because the photographs used were of adult twins, the majority of whom had passed the average age of onset for BN (sample mean age 30.1 years). Thus, it is unclear whether physical similarity was evident before the development of BN, or whether concordant twins became more physically similar because of the effects of eating disorder symptomatology

https://doi.org/10.1017/S003329170999122X Published online by Cambridge University Press

Table 1. *Item numbers, corresponding interview questions, and scoring*

Item no.	Interview question	Possible response (frequency)
1[a]	Have you ever had eating binges when you ate what most people would describeas an unusually large amount of food in a short time?	(0) No (817) (1) Yes (207)
2	When you were having eating binges, did you feel that your eating was out of control?	(0) Not at all (33) (1) Slightly (26) (2) Somewhat (65) (3) Very much (61) (4) Extremely (44)
3	When you were bingeing the most, how many binges would you have in a month?	(0) 0–1 (44) (1) 2–7 (99) (2) ⩾8 (64)
4	For how long did you have binge eating episodes?	(0) Less than 1 month (62) (1) 1 to 2 months (19) (2) 3 months to 5 months (14) (3) 6 months to 1 year (40) (4) longer than 1 year (68)
Items 5–11 began with the stem: 'During your most extreme efforts to control your shape and weight, how often did you do the following:'		
5	Make yourself vomit	(0) Never (918) (1) Once (27) (2) Less than once per week (12) (3) Once per week (7) (4) A few days per week (14) (5) Nearly every day (7) (6) Every day (10)
6	Laxatives	(0) Never (862) (1) Once (41) (2) Less than once per week (41) (3) Once per week (14) (4) A few days per week (19) (5) Nearly every day (11) (6) Every day (5)
7	Diuretics (water pills)	(0) Never (853) (1) Once (48) (2) Less than once per week (43) (3) Once per week (11) (4) A few days per week (17) (5) Nearly every day (12) (6) Every day (9)
8	Diet pills (over the counter or prescription)	(0) Never (580) (1) Once (146) (2) Less than once per week (44) (3) Once per week (21) (4) A few days per week (69) (5) Nearly every day (60) (6) Every day (72)
9	Exercise more than 2 h per day	(0) Never (760) (1) Once (23) (2) Less than once per week (56) (3) Once per week (31) (4) A few days per week (73) (5) Nearly every day (33) (6) Every day (15)

https://doi.org/10.1017/S003329170999122X Published online by Cambridge University Press

Table 1 (*cont.*)

Item no.	Interview question	Possible response (frequency)
10	Fast or not eat (for 24 h or more)	(0) Never (811) (1) Once (67) (2) Less than once per week (47) (3) Once per week (22) (4) A few days per week (27) (5) Nearly every day (7) (6) Every day (4)
11	Choose one of the following statements that best describes you[b]	(0) Weight or shape is not at all important to how I feel about myself (29) (1) Weight or shape plays a small part in how I feel about myself (167) (2) Weight or shape plays a moderate part in how I feel about myself (425) (3) Weight or shape plays a major part in how I feel about myself (335) (4) Weight or shape is the most important thing that affects how I feel about myself (56)

[a] If participants indicated that they had never binged, they skipped to the next section of the questionnaire.
[b] This is a general question and is not specific to any time period.

on their appearance. Consequently, Klump *et al*. (2000*a*) studied adolescents and included multiple measures of physical similarity, including photographs, body shape ratings and BMI. They found no associations between physical similarity and twins' scores on a measure of eating attitudes and behaviors, supporting the validity of the EEA.

The current study incorporates factor scores on latent EEA traits into a twin model as definition variables to determine the magnitude of potential EEA violations. If results indicate that the EEA is violated, then those aspects that contribute to twin similarity in eating pathology can be used as moderators in subsequent research. Further, this study uses a recently validated measure of the EEA (Mitchell *et al*. 2007), which assesses both childhood and adolescent aspects of twins' environment. Measurement of both these time periods is particularly relevant to studies of eating disorders, as genetic and environmental influences on liability appear to operate differently in girls of different ages (e.g. Klump *et al*. 2000*b*, 2007; Silberg & Bulik, 2005).

Method

Participants

This study includes MZ ($n=614$) and DZ ($n=410$) female twins from the population-based Virginia Twin Registry (Kendler & Prescott, 1999), now part of the Mid-Atlantic Twin Registry. DZ female twins with male co-twins were not included in these analyses. A total of 27 participants met criteria for BN, as defined by DSM-IV criteria (APA, 2000). The mean age of twins in this study was 40.44 (S.D. = 8.34).

These data, which were part of a multi-wave, ongoing data collection, have been described elsewhere (Kendler & Prescott, 2006). Of note, this sample overlaps with those used in earlier studies of BN (Kendler *et al*. 1991, 1995; Bulik *et al*. 1998).

Measures

Bulimia nervosa symptoms

Items based on DSM-IV (APA, 2000) criteria for BN were adapted from the Structured Clinical Interview for DSM-IV (First *et al*. 1997) and administered in a self-report survey (see Table 1 for a summary of items). Of note, if participants reported never binge eating, they skipped subsequent items related to binge eating; however, they did answer items assessing compensatory behaviors and the influence of weight on self-evaluation. A total of 207 women reported that they had ever binged. Binge eating was scored dichotomously; all other items were assessed using ordinal scales. The item assessing the number of binges per month used a free-response format. These responses were trichotomized (0–1, 2–7 and $\geqslant 8$ binges per month) for model-fitting

https://doi.org/10.1017/S003329170999122X Published online by Cambridge University Press

Equal environments

Seven items measuring twins' equal environments were included in the survey (Kendler & Gardner, 1998). These items, which assess childhood and teenage environmental similarity, were based on Loehlin & Nichols' (1976) work. Using exploratory and confirmatory factor analytic techniques, Mitchell *et al*. (2007) extracted two correlated ($r = 0.59$) factors, labeled 'child' and 'teen' and confirmed that two items loaded on each factor. The child factor is assessed by the items, 'when you were children, up to the age of 13, how often did you and your twin share the same room?' and 'when you were children, how often did you and your twin dress alike?' The teen factor is assessed by the items, 'as teenagers, how often would you and your twin have the same friends?' and 'as teenagers, how often would your twin go out with you if you went to the movies or a dance?' In the current study, factor scores were created for the child and teen factors; these were used as definition variables in the model of the impact of equal environments on BN. Note, only one factor score per twin pair was used in the models.

Analyses

Factor scores were estimated from the factor model of the equal environments items using Mplus (Muthén & Muthén, 1998–2006). Mx (Neale *et al*. 2003) was used for twin modeling. An item-factor modeling approach (Neale *et al*. 2006*a*) was used for the analyses. This method, which has been applied to the analysis of AN symptoms (Mazzeo *et al*. 2009), is a latent trait model that is formally equivalent to a two-parameter normal ogive item response and is an application of the common factor model to multivariate binary or ordinal data, such that the likelihood of item data is computed conditional on the latent trait. The model estimates a 'location' on the factor liability scale for each diagnostic criterion, which is the point on the liability scale where there is 0.5 probability of endorsing the criterion. That portion of variance unique to each BN criterion, and not accounted for by the common factor, was also estimated. Marginal maximum likelihood estimation was used, in which the overall likelihood is computed by integrating over the latent trait. This method utilizes a finite mixture distribution, which is specified for points on the latent trait. Gaussian quadrature weights were used to weight the likelihood at latent trait values; the weighted likelihoods are summed in order to compute the overall likelihood. Of note, use of at least 10 points provides a good approximation (Neale *et al*. 2006*a*).

Due to skip patterns in the survey (regarding binge eating, noted above), there were missing data. Selection effects from 'gateway' items can impact estimation of the covariation among the items, which is essential for fitting the factor model. Specifically, for a single binary gateway item with individual record data, there will be no variance on the gateway item when data on the probe items are available, because individuals must endorse the gateway item in order to be asked the probe items. Ultimately, this zero variance problem can affect validity of factor analyses. However, joint analysis of gateway and probe items collected from twin pairs overcomes this problem, because the covariance between the gateway item and the co-twin's probe items is available (Neale *et al*. 2006*b*).

Three main types of parameters are estimated in this model: (1) thresholds (location), which reflect the probabilities that the BN symptoms are endorsed; (2) factor loadings, which estimate association between the latent trait and each of the symptoms; (3) sources of variance of the latent trait. The factor variance is partitioned into influences due to additive genetic (A), shared environment (C) and specific or individual environment (E) effects. A fourth source of variance was also estimated: that due to shared twin experiences (M). Factor scores (first child and then teen) were added as a moderator of the shared environment covariance parameter (Fig. 1).

In biometric modeling, additive genetic effects are specified to contribute twice as much to the covariance between MZ twins as DZ twins because, generally, MZ twins share all of their genes, whereas DZ twins on average share half. Environmental influences shared by members of a twin pair are specified to correlate 1.0 among twins, regardless of whether they are MZ or DZ. Specific environmental influences are those that are assumed to be unique to each twin and therefore uncorrelated among MZ and DZ twin pairs. In this model, each item's observed variance is decomposed into two types: that shared with the latent factor and that portion unique to the item (residual or unique variance), including measurement error. Residual variance for each item was partitioned into A, C, and E influences.

An ACE model was compared with the full ACEM model to test for EEA violations using a likelihood ratio test ($\Delta\chi^2$). A significant χ^2 difference indicates that model fit worsens when parameters are fixed to zero. If dropping the path from M does not significantly worsen model fit, the EEA is tenable. CIs are reported as well, to evaluate further the statistical precision of the measured similarity parameter and of the A, C and E parameters. Akaike's Information Criterion (AIC) values for the models, computed as -2lnL to 2df (Akaike, 1987) are also reported. However, this index, under certain conditions, may

https://doi.org/10.1017/S003329170999122X Published online by Cambridge University Press

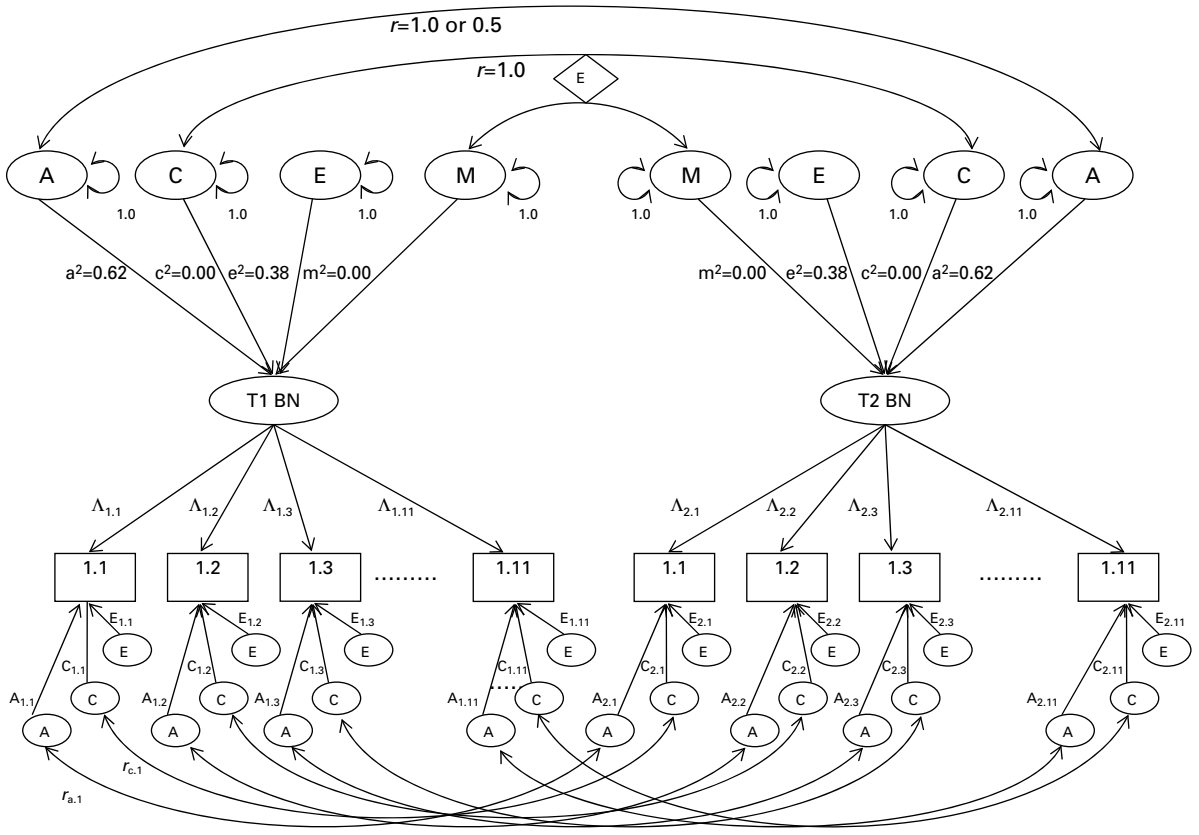

Fig. 1. Factor scores. Only four items per twin are shown. ◇ = factor scores on the latent equal environment assumption trait as a definition variable. $\Lambda_{1.1}$ denotes the factor loading for the first item, for twin 1. $A_{1.1}$ denotes the residual variance due to A for the first item, for twin 1. $r_{a.1}=1.0$ or 0.5 and denotes the twin correlation for the residual components due to A for the first item, etc. $r_{c.1}=1.0$ and denotes the twin correlation for the residual components due to C for the first item, etc.

over-simplify models (Sullivan & Eaves, 2002) and was not used to evaluate model fit.

Results

Child equal environment factor

An ACEM item-factor model was first fit to the MZ and DZ twin data (see Table 2 for fit statistics for all models). The MZ correlation ($r=0.62$) was higher than the DZ correlation ($r=0.31$), suggesting that BN is heritable. The largest proportion of variance in latent BN liability was due to additive genetic effects ($a^2=0.62$, 95% CI 0.18–0.76), with the remainder due to unique environmental influences ($e^2=0.38$, 95% CI 0.23–0.62); the effects of the common environment ($c^2=0.00$, 95% CI 0.00–0.35) and measured similarity ($m^2=0.00$, 95% CI 0.00–0.001) were estimated at nearly zero. A comparison of the ACE model to the full ACEM model indicated that dropping the moderation parameter on the shared environment covariance did not significantly worsen model fit, as indicated by a χ^2 test ($\Delta\chi^2=0.003$, df = 1, $p>0.05$).

Thus, the AE model provided the most parsimonious fit (AIC −3708.788).

This model also provides information regarding the covariation among the BN criteria and their variances. This information is obtained by examination of the latent BN factor loadings and residual variances. All factor loadings were significant, ranging from 0.45 to 0.82 (Table 3). Residual variance for each item was partitioned into A, C and E influences. For nearly all items, the largest amount of residual variance was due to unique environment and measurement error (Table 3). However, the majority of residual variance for vomiting (40%; item 5) and diet pills (43%; item 8), which are compensatory behaviors, was due to additive genetic effects. Several items had modest residual variance due to common environmental factors: binge eating (21%, item 1); duration of binges (27%, item 4); use of vomiting (32%, item 5); excessive exercise (17%, item 9); fasting (20%, item 10); influence of weight and shape on self evaluation (69%, item 11).

Total heritability for each individual item (i) was computed as the product of the item's squared factor loading (λ) and a^2 for the latent trait, added to the

https://doi.org/10.1017/S003329170999122X Published online by Cambridge University Press

Table 2. *Summary of fit information for the full ACEM model as well as ACE, AE, and CE submodels*

Model	−2LL	df	$\Delta\chi^2$	Δdf	*p*	AIC	A^2 (CI)	C^2 (CI)	E^2 (CI)	M^2 (CI)
Child similarity factor										
Full ACEM	13270.780	8485	–	–	–	−3699.22	0.62 (0.18–0.76)	0.00 (0.00–0.35)	0.38 (0.23–0.62)	0.00 (0.00–0.001)
ACE	13270.777	8486	0.003	1	>0.05	−3701.223	–	–	–	–
Teen similarity factor										
Full ACEM	13265.808	8485	–	–	–	−3704.192	0.63 (0.26–0.74)	0.00 (0.00–0.00)	0.37 (0.24–0.56)	0.02 (0.00–0.11)
ACE	13267.540	8486	1.73	1	>0.05	−3704.460	–	–	–	–

−2LL, −2 log-likelihood; AIC, Akaike's Information Criterion; CI, 95% confidence intervals; A, additive genetic influence on the latent trait; C, common environmental influence on the latent trait; E, unique environmental influence on the latent trait.

Table 3. *Item factor loadings, residual variances, and heritability estimates* (*95% confidence intervals*)

Item	Factor loading	Residual variance			Total heritability	Total common environment	Total unique environment
		A	C	E			
1. Ever had eating binges	0.65 (0.51–0.76)	0.01	0.21	0.65	0.34 (0.21–0.53)	0.12 (0.00–0.31)	0.54 (0.39–0.74)
2. Eating out of control during binges	0.74 (0.56–0.82)	0.12	0.09	0.79	0.39 (0.25–0.71)	0.04 (0.00–0.18)	0.57 (0.26–0.75)
3. Frequency of binges per month	0.75 (0.51–0.88)	0.17	0.12	0.71	0.41 (0.23–0.67)	0.06 (0.00–0.22)	0.53 (0.25–0.74)
4. Duration of binge eating episodes	0.55 (0.23–0.74)	0.16	0.27	0.58	0.30 (0.11–0.71)	0.19 (0.00–0.64)	0.52 (0.13–0.80)
5. Vomiting	0.75 (0.59–0.85)	0.40	0.32	0.28	0.53 (0.41–0.74)	0.14 (0.00–0.24)	0.33 (0.22–0.48)
6. Laxatives	0.82 (0.71–0.88)	0.03	0.06	0.91	0.43 (0.21–0.54)	0.02 (0.00–0.12)	0.55 (0.39–0.79)
7. Diuretics	0.67 (0.53–0.78)	0.30	0.07	0.65	0.44 (0.23–0.62)	0.03 (0.00–0.22)	0.53 (0.36–0.74)
8. Diet pills	0.66 (0.55–0.72)	0.43	0.00	0.57	0.51 (0.23–0.62)	0.00 (0.00–0.09)	0.49 (0.38–0.77)
9. Excessive exercise	0.51 (0.36–0.61)	0.25	0.17	0.58	0.35 (0.13–0.60)	0.12 (0.00–0.32)	0.53 (0.38–0.70)
10. Fasting	0.68 (0.54–0.77)	0.25	0.20	0.55	0.42 (0.20–0.62)	0.11 (0.00–0.27)	0.47 (0.34–0.65)
11. Importance of weight/ shape to self-evaluation	0.44 (0.32–0.55)	0.14	0.69	0.69	0.24 (0.10–0.46)	0.13 (0.00–0.28)	0.63 (0.53–0.75)

A, Additive genetic influence on the latent trait; C, common environmental influence on the latent trait; E, unique environmental influence on the latent trait.

product of one minus the item's squared factor loading and the amount of the item's residual variance due to A, termed as^2. This equation, where λ_i is the factor loading for the *i*th item, is as follows:

$$(\lambda_i^2)(a^2)+(1-\lambda_i^2)(as_i^2).$$

Similarly, total shared and unique environmental influences on each item were computed using this equation, respectively substituting c^2 or e^2 and residual variance due to C or E. Four items (5, 6, 7, 8), assessing use of vomiting, laxatives, diuretics and diet pills as compensatory behaviors, had heritability estimates ranging from 0.43 to 0.53. Loss of control during binges (item 2), frequency of binge eating (item 3) and fasting as compensatory behavior (item 10) had estimates ranging from 0.39 to 0.42. Heritability estimates for binge eating (item 1) and use of excessive exercise as a compensatory behavior (item 9) were 0.34 and 0.35, respectively. Finally, binge eating duration (item 4) and the effect of weight and shape on self-evaluation (item 11) had heritability estimates of 0.30 and 0.24, respectively.

Teen equal environment factor

The same model was fit to the data using factor scores on the teen similarity factor as a definition variable.

https://doi.org/10.1017/S003329170999122X Published online by Cambridge University Press

We do not present factor loadings and item-specific variances for this model, as these results are nearly identical to the previous model. As with the child factor, the effect of measured similarity was estimated at nearly zero (95% CI 0.00–0.11). Dropping this parameter from the full ACEM model did not significantly worsen fit (see Table 2).

Discussion

Diagnosis and classification of eating disorders, including BN, has recently received a great deal of scrutiny (e.g. Bulik *et al*. 2007) and some have suggested that current DSM-IV criteria have 'fundamental flaws' (Fairburn & Cooper, 2007, p. S107). One challenge is that current eating disorder diagnoses comprise a range of specific symptoms, of which a minimum count is required to meet diagnostic criteria. This may not be optimal if individual symptoms do not relate to the underlying disorder phenotype in the same way. Further, from a genetic epidemiological perspective, it is plausible that distinct criteria are differentially influenced by genetic or environmental factors. This study used an item-factor model to investigate the influence of genetic and environmental factors to both the overall BN diagnosis and to specific symptoms. This study also evaluated the validity of the EEA, applied to BN, as this assumption is fundamental to twin research.

Results indicated that BN liability was significantly influenced by additive genetic factors; unique environmental factors also significantly influenced BN at the factor level, but to a lesser degree. In a study examining the genetic epidemiology of broadly defined BN, Bulik *et al*. (1998) also found that an AE model provided the best fit. Further, these authors found a similar heritability estimate (60%) when BN was assessed using a single interview.

Item-level analyses suggest that, although the overall diagnosis of BN may be highly influenced by genetic factors, not all symptoms are equally heritable. In particular, current results are consistent with studies supporting the role of environmental factors on the influence of weight on self-evaluation (e.g. Wade *et al*. 1998; Reichborn-Kjennerud *et al*. 2004). Further, the unique environment most strongly influenced liability to this symptom in this study. Thus, perhaps, important targets for intervention and prevention are experiences that occur outside of the family, such as competing in an appearance-oriented activity or sport (e.g. dance, modelling, or gymnastics), teasing by peers or stressors such as abuse, which are only experienced by one member of a twin pair. However, additional research is needed to evaluate this hypothesis, because (E), the component of the model that evaluates unique environmental experiences, also assesses variance attributable to measurement error. These two sources of variance cannot be separated in traditional twin models. Thus, future studies should attempt to extend the current findings and evaluate measurement error and unshared experiences separately.

Results of this study also indicated that self-induced vomiting was the symptom most strongly influenced by additive genetic factors. Sullivan *et al*. (1998) also found a strong influence of additive genetic effects on self-induced vomiting, although their heritability estimate was somewhat higher. Of note, binge eating was only moderately influenced by additive genetic effects; the unique environment contributed most strongly to variance in this symptom. This heritability estimate is lower than that obtained in previous studies of binge eating (which have ranged from 41% to 82%; Bulik *et al*. 1998, 2007; Sullivan *et al*. 1998; Reichborn-Kjennerud *et al*. 2003). However, differences in analytic strategy, particularly the current study's use of the item-factor approach, might partially account for these inconsistencies.

The remainder of BN symptoms assessed in this study appear to be best accounted for by a mixture of moderately strong additive genetic factors and unique environmental influences. As noted above, the unique environment component of item-level variance in this model also includes measurement error. Repeated measures might be used to estimate variance due to this latter source. These results are consistent with those of another investigation (Wade *et al*. 2006), which found that unique environmental factors contributed substantially to the number of lifetime eating disordered behaviors. Future research should examine the influence of specific experiences on the development of eating disorder symptoms in individuals with known genetic risk for AN or BN. Such work would facilitate understanding of potential gene environment interactions and inform prevention. Individual differences in methylation may also contribute to the specific environment variance component, as MZ twins do not correlate perfectly for their gene activation (Kato *et al*. 2005). In addition, further investigation of the structure of BN is warranted; the current results suggest that, although all items had relatively large loadings on the latent diagnostic factor, some, most notably the influence of weight on self-evaluation, were lower.

Finally, results of the current study suggested that the EEA was not violated. It is important to examine this assumption because it is fundamental to the validity of results yielded by twin modeling. This study used a recently validated measure of the EEA (Mitchell *et al*. 2007) and found no evidence of any

https://doi.org/10.1017/S003329170999122X Published online by Cambridge University Press

violations, consistent with the majority of previous studies in this area (e.g. Kendler *et al*. 1993; Klump *et al*. 2000*a*).

Limitations of this study should be noted. First, the sample included exclusively European–American female twins. It is unclear whether these results are applicable to men, non-twins or other cultural groups. Second, like many other twin studies, our sample includes a greater proportion of MZ twins and individuals with above-average levels of education (Lykken *et al*. 1987). Third, BN symptoms were assessed using single-item measures, which might attenuate reliability (Crocker & Algina, 1986) and potentially confound heritability estimates. However, given that this was a large population-based survey that addressed multiple disorders and their correlates, the survey needed to be as concise as possible to reduce participant burden and fatigue. Finally, lifetime BN symptoms were evaluated in this study to increase statistical power. Thus, results may have been influenced by recall bias.

Despite these limitations, this study has several strengths, including the use of a large, population-based sample and symptom level modeling. Such results provide much richer data informative to the development of liability indices (Bulik *et al*. 2007). Future studies should extend this work to inform treatment and prevention efforts.

Acknowledgements

This research was supported by the National Institutes of Health Grants MH-068520 (Mazzeo), MH-20030 (Mitchell), MH66117–05 (Bulik, Devlin PI), MH-65322 (Aggen and Neale, Neale PI), and MH-40828 (Kendler). We are very grateful to the twins for their participation.

Declaration of interest

None.

References

Akaike H (1987). Factor analysis and AIC. *Psychometrika* **52**, 317–332.

APA (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn, text revision. American Psychiatric Association: Washington, DC.

Becker AE, Fay K (2006). Sociocultural issues and eating disorders. In *Eating Disorders Review Part II* (ed. S. Wonderlich, J. E. Mitchell, M. de Zwaan and H. Steiger), pp. 35–63. Radcliffe Publishing: Oxford.

Becker AE, Grinspoon SK, Klibanski A, Herzog DB (1999). Eating disorders. *New England Journal of Medicine* **340**, 1092–1098.

Bulik CM (2005). Exploring the gene-environment nexus in eating disorders. *Journal of Psychiatry and Neuroscience* **30**, 335–338.

Bulik CM, Hebebrand J, Keski-Rahkonen A, Klump KL, Reichborn-Kjennerud T, Mazzeo SE, Wade TD (2007). Genetic epidemiology, endophenotypes, and eating disorder classification. *International Journal of Eating Disorders* **40**, S52–S60.

Bulik CM, Sullivan PF, Kendler KS (1998). Heritability of binge-eating and broadly defined bulimia nervosa. *Biological Psychiatry* **44**, 1210–1218.

Bulik CM, Sullivan PF, Wade TD, Kendler KS (2000). Twin studies of eating disorders: a review. *International Journal of Eating Disorders* **27**, 1–20.

Cachelin FM, Veisel C, Barzegarnazari E, Striegel-Moore, RH (2000). Disordered eating, acculturation, and treatment-seeking in a community sample of Hispanic, Asian, Black, and White women. *Psychology of Women Quarterly* **24**, 244–253.

Crocker L, Algina J (1986). *Introduction to Classical and Modern Test Theory*. Wadsworth Group: Belmont, CA.

Fairburn CG, Cowen PJ, Harrison PJ (1999). Twin studies and the etiology of eating disorders. *International Journal of Eating Disorders* **26**, 349–358.

Fairburn CG, Cooper Z (2007). Thinking afresh about the classification of eating disorders. *International Journal of Eating Disorders* **40**, S107–S110.

First MB, Spitzer RL, Gibbon M, Williams JBW (1997). *Structured Clinical Interview for DSM-IV Axis I Disorders* (*SCID-I*), *Clinician Version*. American Psychiatric Press: Washington, DC.

Hettema JM, Neale MC, Kendler KS (1995). Physical similarity and the equal-environment assumption in twin studies of psychiatric disorders. *Behavior Genetics* **25**, 327–335.

Hudson JI, Hiripi E, Pope HG, Kessler RC (2007). The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. *Biological Psychiatry* **61**, 348–358.

Kato T, Iwamoto K, Kakiuchi C, Kuratomi G, Okazaki Y. (2005). Genetic or epigenetic difference causing discordance between monozygotic twins as a clue to molecular basis of mental disorders. *Molecular Psychiatry* **10**, 622–630.

Kendler KS, Gardner CO (1998). Twin studies of adult psychiatric and substance dependence disorders: Are they biased by differences in the environmental experiences of monozygotic and dizygotic twins in childhood and adolescence? *Psychological Medicine* **28**, 625–633.

Kendler KS, MacLean CJ, Neale MC, Kessler RC, Heath AC, Eaves LJ (1991). The genetic epidemiology of bulimia nervosa. *American Journal of Psychiatry* **148**, 1627–1637.

Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ (1993). A test of the equal environment assumption in twin studies of psychiatric illness. *Behavior Genetics* **23**, 21–27.

Kendler KS, Prescott C (1999). A population-based twin study of lifetime major depression in men and women. *Archives of General Psychiatry* **56**, 39–44.

Kendler KS, Prescott C (2006). *Genes, Environment, and Psychopathology: Understanding the Causes of Psychiatric*

and Substance Use Disorders. The Guilford Press: New York, NY.

Kendler KS, Walters EE, Neale MC, Kessler RC, Heath AC, Eaves LJ (1995). The structure of the genetic and environmental risk factors for six major psychiatric disorders in women: phobia, generalized anxiety disorder, panic disorder, bulimia, major depression and alcoholism. *Archives of General Psychiatry* **52**, 374–383.

Keski-Rahkonen A, Bulik CM, Neale BM, Rose RJ, Rissanen A, Kaprio J (2005). Body dissatisfaction and drive for thinness in young adult twins. *International Journal of Eating Disorders* **37**, 188–199.

Klump KL, Burt SA, McGue M, Iacono WG (2007). Changes in genetic and environmental influences on disordered eating across adolescence: a longitudinal twin study. *Archives of General Psychiatry* **64**, 1409–1415.

Klump KL, Holly A, Iacono WG, McGue M, Willson LE (2000*a*). Physical similarity and twin resemblance for eating attitudes and behaviors: a test of the equal environments assumption. *Behavior Genetics* **30**, 51–58.

Klump KL, McGue M, Iacono WG (2000*b*). Age differences in genetic and environmental influences on eating attitudes and behaviors in preadolescent and adolescent female twins. *Journal of Abnormal Psychology* **109**, 239–251.

le Grange D (2005). Family issues and eating disorders. In *Eating Disorders Review Part I* (ed. S. Wonderlich, J. E. Mitchell, M. De Zwaan, and H. Steiger), pp. 15–25. Radcliffe Publishing: Oxford.

Loehlin JC, Nichols RC (1976). *Heredity, Environment, and Personality: A Study of 850 Sets of Twins*. University of Texas Press: Austin, TX.

Lykken DT, McGue M, Tellegen A (1987). Recruitment bias in twin research: The rule of two-thirds reconsidered. *Behavior Genetics* **17**, 343–362.

Mazzeo SE, Mitchell KS, Bulik CM, Reichborn-Kjennerud T, Kendler KS, Neale MC (2009). Assessing the heritability of anorexia nervosa symptoms using a marginal maximal likelihood approach. *Psychological Medicine* **39**, 463–473.

Mitchell KS, Mazzeo SE, Bulik CM, Aggen SH, Kendler KS, Neale MC (2007). An investigation of a measure of twins' equal environments. *Twin Research and Human Genetics* **10**, 840–847.

Muthén LK, Muthén BO (1998–2006). *Mplus User's Guide*, 4th edn. Muthén & Muthén: Los Angeles, CA.

Neale MC, Aggen SH, Maes HH, Kubarych TS, Schmitt JE (2006*a*). Methodological issues in the assessment of substance use phenotypes. *Addictive Behaviors* **31**, 1010–1034.

Neale MC, Boker S, Xie G, Maes H (2003). *Mx: Statistical modeling*, 6th edn. Department of Psychiatry, Virginia Commonwealth University: Richmond, VA.

Neale MC, Harvey E, Maes H, Sullivan PF, Kendler KS (2006*b*). Extensions to the modeling of initiation and progression: Applications to substance use and abuse. *Behavior Genetics* **36**, 507–524.

Neale MC, Lubke G, Aggen SH, Dolan CV (2005). Problems with using sum scores for estimating variance components: Contamination and measurement noninvariance. *Twin Research and Human Genetics* **8**, 553–568.

Reichborn-Kjennerud T, Bulik CM, Kendler KS, Røysamb E, Maes H, Tambs K, Harris J (2003). Gender differences in binge-eating: a population-based twin study. *Acta Psychiatrica Scandinavica* **108**, 196–202.

Reichborn-Kjennerud T, Bulik CM, Kendler KS, Roysamb E, Tambs K, Torgersen S, Harris JR (2004). Undue influence of weight on self-evaluation: a population-based twin study of gender differences. *International Journal of Eating Disorders* **35**, 123–132.

Silberg JL, Bulik CM (2005). The developmental association between eating disorders symptoms and symptoms of depression and anxiety in juvenile twin girls. *Journal of Child Psychology and Psychiatry* **46**, 1317–1326.

Simon S, Schmidt U, Pilling S (2005). The health service use and cost of eating disorders. *Psychological Medicine* **35**, 1543–1551.

Spoor STP, Stice E, Burton E, Bohon C (2007). Relations of bulimic symptom frequency and intensity to psychosocial impairment and health care utilization: Results from a community-recruited sample. *International Journal of Eating Disorders* **40**, 505–514.

Sullivan PF, Bulik CM, Kendler KS (1998). Genetic epidemiology of binging and vomiting. *British Journal of Psychiatry* **173**, 75–79.

Sullivan PF, Eaves LJ (2002). Evaluation of analyses of univariate discrete twin data. *Behavior Genetics* **32**, 221–227.

Wade TD, Bergin JL, Martin NG, Gillespie NA, Fairburn CG (2006). A transdiagnostic approach to understanding eating disorders. *The Journal of Nervous and Mental Disease* **194**, 510–517.

Wade TD, Martin NG, Tiggemann M (1998). Genetic and environmental risk factors for the weight and shape concerns characteristic of bulimia nervosa. *Psychological Medicine* **28**, 761–771.

Wilfley DE, Bishop ME, Wilson GT, Agras WS (2007). Classification of eating disorders: Toward DSM-V. *International Journal of Eating Disorders* **40**, S123–S129.

Winn S, Perkins S, Walwyn R, Schmidt U, Eisler I, Treasure J, Berelowitz M, Dodge L, Frost S, Jenkins M, Johnson-Sabine E, Keville S, Murphy R, Robinson P, Yi I (2007). Predictors of mental health problems and negative caregiving experiences in carers of adolescents with bulimia nervosa. *International Journal of Eating Disorders* **40**, 171–178.