Behav Genet (2013) 43:427–435
DOI 10.1007/s10519-013-9603-0

ORIGINAL RESEARCH

# An Examination of the Representativeness Assumption for Twin Studies of Eating Pathology and Internalizing Symptoms

**Melissa A. Munn-Chernoff · Kristin M. von Ranson · Kristen M. Culbert · Christine L. Larson · S. Alexandra Burt · Kelly L. Klump**

Received: 20 December 2012 / Accepted: 19 July 2013 / Published online: 30 July 2013


**Abstract** Little research has investigated whether the twin representativeness assumption (that results from twin research generalize to singletons) holds for eating pathology and internalizing symptoms. This study compared disordered eating, depression, and anxiety among young adult female twins versus singletons. Participants included 292 twins and 997 singletons in three samples. Questionnaires included the Minnesota Eating Behavior Survey, Eating Disorder Examination Questionnaire, Beck Depression Inventory, and State-Trait Anxiety Inventory. We examined mean differences between twins' and singletons' scores, after adjusting for age, body mass index, and ethnicity. We found statistically significant mean differences on psychopathology, with twins reporting less disordered eating and internalizing symptoms compared with singletons. Effect sizes of these mean differences were small to moderate. Our results suggest that twins report less disordered eating and internalizing symptoms than singletons, which, combined with the generally small effect sizes, indicate that results from twin samples generalize to singletons.

**Keywords** Disordered eating · Depression · Anxiety · Twin representativeness assumption · Singletons

Edited by Valerie Knopik.

M. A. Munn-Chernoff
Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, USA

M. A. Munn-Chernoff
Midwest Alcoholism Research Center, Washington University School of Medicine, St. Louis, MO, USA

K. M. von Ranson (✉)
Department of Psychology, University of Calgary, 2500 University Drive NW, Calgary, AB T2N 1N4, Canada
e-mail: kvonrans@ucalgary.ca

K. M. Culbert · S. A. Burt · K. L. Klump
Department of Psychology, Michigan State University, East Lansing, MI, USA

C. L. Larson
Department of Psychology, University of Wisconsin, Milwaukee, WI, USA

## Introduction

Over the past few decades, twin research has provided much information about the genetic and environmental architecture of psychopathology. In adulthood, studies have suggested that genetic influences are present for many traits—including eating disorders and their symptoms (Thornton et al. 2011), major depression (Kendler et al. 2006), and anxiety (Mosing et al. 2009)—with non-shared environmental influences generally explaining the remaining portion of variance affecting each trait. Despite the progress that has been made disentangling the relative contribution of genetic and environmental influences on these traits, few studies have examined a critical assumption of twin studies, namely that twins are representative of the general population.

The twin representativeness assumption states that twins are representative of the general population for a given trait of interest. Thus, this assumption holds that twins do not have different rates of psychopathology compared with non-twin (i.e., singleton) individuals. Critics of twin methodology argue that twin research is not generalizable to singleton research because twins are often born

premature, have lower birth weights, and are at increased risk for perinatal complications and perinatal death compared with singletons, all of which could conceivably affect rates of psychopathology (Evans and Martin 2000). One can test the representativeness assumption for rates of psychopathology by investigating whether the two groups significantly differ from each other on constructs of interest. If twins and singletons are significantly different for a given trait, the assumption may be violated and findings from twin research may not generalize to singletons (Kendler et al. 1995; Klump et al. 1999).

Although it has been argued that twins have higher rates of psychopathology compared with singletons (Gau et al. 1992; Gjone and Novik 1995), thus violating the twin representativeness assumption, prior studies have been conflicting. In a review by Kendler (Kendler 1993), rates of psychopathology, which included "functional psychosis", schizophrenia, and manic-depressive disorder, were similar between twin and singleton populations. Likewise, findings from twins generalized to singletons for depression, social anxiety, and generalized anxiety disorder (Kendler et al. 1995, 1996; Pulkkinen et al. 2003). However, twins had higher rates of panic-phobia (i.e., a factor score that included items such as "terror spells", "suddenly scared", "avoiding frightening things", and "nervous if alone") (Kendler et al. 1995) and disordered eating (Klump et al. 1999) in other studies, suggesting variability in the extent to which the representativeness assumption is upheld for all forms of psychopathology. Thus, the limited amount of research in this area, and the somewhat discrepant results, highlight the need for additional research to understand the extent to which twins represent the general population for disordered eating and internalizing symptoms. Validating the twin representativeness assumption is critical because twin and family designs provide a unique resource (e.g., matched controls for environmental factors shared by family members (Hopper et al. 2005)) for elucidating the role of genetic and environmental influences on these symptoms. If the twin representativeness assumption is not upheld, then findings may not generalize to non-twins, which would significantly limit the information that can be gleaned from these studies.

The aim of the current study is to examine whether there is support for the representativeness assumption of twin studies for disordered eating and internalizing symptoms among young adult women within three separate singleton samples. Based on prior research regarding psychopathology in general, we hypothesize that there are no significant mean or variance differences for these traits between twins and singletons. If there are no significant differences between twins and singletons, it is likely that findings from twin research generalize to singleton populations.

## Methods

### Participants

The studies from which participants were drawn were each approved by an institutional review board, and all participants completed informed consent prior to participating.

#### *Twin sample*

Two hundred and ninety-two same-sex young adult female twins (162 monozygotic, 130 dizygotic) were included in these analyses. All twins were part of the Adult Twin Study of Behavioral Adjustment and Development through the Michigan State University Twin Registry, which assessed genetic, environmental, and neurobiological influences on disordered eating, depression, anxiety, attention problems, and substance use. Twins were recruited through a university-based twin registry ($n = 148$), flyers ($n = 84$), a state department of community health ($n = 26$), word of mouth ($n = 16$), paid advertisements ($n = 8$), and twin list-servs ($n = 6$). The recruitment method for four individuals was not available. Although the majority of twins were recruited via a university-based twin registry, there were no mean differences in body mass index (BMI) or any of the disordered eating, depression, or anxiety measures between twins who had at least some college education ($n = 219$) versus those who did not/had not yet received a high school degree ($n = 69$; all $p$'s $> 0.12$). In addition, there were no differences in ethnicity between twins who had at least some college education compared with those who did not have some college education ($\chi^2 = 2.16$, df $= 2$, $p = 0.34$). Zygosity was determined via physical similarity questionnaires, which have greater than 95 % accuracy (Iacono et al. 1999; Lykken et al. 1990; Peeters et al. 1998). More detailed information on the twin sample can be found elsewhere (Burt and Klump 2013; Culbert et al. 2008; Klump and Burt 2006).

#### *Singleton sample*

Overall, 997 female singletons were included in these analyses. Similar to the twins, singletons were primarily recruited from a university. Singletons came from three different studies from the same geographic area, and unless otherwise noted, were recruited through a university department of psychology subject pool. The first sample (singleton sample 1) included 447 women who were recruited in 2002 for a study focused on season of birth effects on disordered eating (Munn et al. 2005). A second sample (singleton sample 2) of 417 undergraduate female singletons participated in a study, conducted in 2005–2006, that examined associations between candidate genes and

disordered eating, depression, and anxiety (Racine et al. 2009). The final sample (singleton sample 3) of 106 female singletons was recruited through the same subject pool or undergraduate classes in 2006 and were included in a study that investigated mean levels of disordered eating between females from opposite-sex twin pairs and female singletons to determine whether elevated prenatal testosterone exposure and/or being reared with a brother contributed to eating pathology (Culbert et al. 2008).

## Measures

### *Disordered eating*

For twins and all singletons, disordered eating was assessed with the Minnesota Eating Behavior Survey (MEBS[1]) (Klump et al. 2000; von Ranson et al. 2005) and the Eating Disorder Examination Questionnaire (EDE-Q) (Fairburn and Beglin 1994). The total score for each measure, as well as individual subscales, were examined. The MEBS subscales included body dissatisfaction (which assessed dissatisfaction with the size of various parts of the body), binge eating (which assessed the tendency to engage in episodes of overeating or thoughts of overeating), compensatory behaviors (which assessed the tendency to use or contemplate using self-induced methods of vomiting and/or diuretics), and weight preoccupation (which assessed preoccupation with dieting, weight, and the pursuit of thinness). Subscales for the EDE-Q included restraint (which assessed restraint over eating, avoidance of food, and dietary avoidance), eating concern (which assessed preoccupation with food, eating in secret, and guilt about eating), shape concern (which assessed the desire for a flat stomach, the importance of body shape, and fear of weight gain), and weight concern (which assessed the importance of weight, dissatisfaction with weight, and the desire to lose weight). Both the MEBS and EDE-Q have shown acceptable psychometric properties (Klump et al. 2000; Mond et al. 2004; Peterson et al. 2007; von Ranson et al. 2005). In the current study, the Cronbach's $\alpha$ for the MEBS total score was 0.90 separately in twins and singletons. For the subscales, the $\alpha$'s were generally $\geq$0.75 among twins and among singletons. The exception was for compensatory behavior, where the Cronbach's $\alpha$ was 0.65 in twins and 0.61 in singletons, which is consistent with previous findings (von Ranson et al. 2005). For the EDE-Q Global Score and subscales, all $\alpha$'s were $\geq$0.80 within twins and within singletons.

[1] The Minnesota Eating Behavior Survey (MEBS; previously known as the Minnesota Eating Disorder Inventory (M-EDI)) was adapted and reproduced by special permission of Psychological Assessment Resources, Inc., 16204 North Florida Avenue, Luz, Florida 33549, from the Eating Disorder Inventory (collectively, EDI and EDI-2) by Garner, Olmstead, Polivy. Copyright 1983 by Psychological Assessment Resources, Inc. Further reproduction of the MEBS is prohibited without prior permission from Psychological Assessment Resources, Inc.

### *Depression and anxiety symptoms*

The Beck Depression Inventory-II (BDI-II) (Beck et al. 1996) was utilized to assess depressive symptoms in the twin sample and in singleton sample 1. In singleton sample 2, the original Beck Depression Inventory (BDI) (Beck et al. 1961) was used to assess depressive symptoms. The mean [standard deviation (SD)] BDI-II score in twins was 8.89 (8.54), whereas in singleton sample 1 it was 10.32 (9.26); the mean BDI score in singleton sample 2 was 10.05 (7.38). We standardized the BDI-II and original BDI total scores before combining them into an overall total depression score for all analyses. Prior studies have reported acceptable reliability and validity for the total score (Beck et al. 1988; Dozois et al. 1998). In the current study, the reliability was adequate (twins: $\alpha = 0.92$; singletons: $\alpha$s $\geq$0.84). Anxiety symptoms were measured with the State-Trait Anxiety Inventory (STAI) (Spielberger et al. 1983), which measured temporary (i.e., state) and stable or enduring (i.e., trait) anxiety, among twins and singleton samples 1 and 2. The STAI has also shown acceptable reliability and validity (Spielberger et al. 1983). The internal consistency was excellent for state (twins: $\alpha = 0.92$; singletons: $\alpha = 0.91$) and trait (twins: $\alpha = 0.94$; singletons: $\alpha = 0.92$) anxiety in the current study. Depression and anxiety symptoms were not assessed in singleton sample 3.

### *Demographic variables*

Age and ethnicity were based on self-reports in both twins and singletons. BMI (calculated as weight [in kilograms]/height [in meters] squared) was based on laboratory assessments of height and weight in the twins and self-report of height and weight in the singletons.

## Statistical analyses

All analyses were conducted in SPSS version 20 (IBM 2011). Hierarchical linear models were used to examine mean differences in disordered eating, depression, and anxiety between twins and singletons. These models are ideal to examine the questions asked in this study because the non-independence of the twin data can be accounted for by nesting a level-one variable (i.e., twin identification variable) within a level-two unit (i.e., family identification variable). Two models were conducted: (1) an initial/unadjusted model that examined the participant variable

(i.e., twin vs. singleton) as the only predictor of the dependent variable (e.g., MEBS total score), and (2) a covariate/adjusted model that controlled for age, BMI, and ethnicity. Age and BMI were continuous variables, whereas ethnicity was a three-level variable, where 1 indicated European–American, 2 indicated African–American, and 3 indicated other ethnic groups. Variance differences on measures of disordered eating, depression, and anxiety in twin 1 versus singletons were also examined using Levene's test for equality of variances. Since twins were randomly labeled as 'twin 1' or 'twin 2', we included only twin 1 in the analysis in order to reduce the number of tests.[2]

In order to detect any violations of the twin representativeness assumption, we wanted to be liberal in our approach to examining sample differences. We therefore reported raw $p$-values in the tables in order to reveal any possible group differences. However, because a large number of comparisons were performed, we also used Bonferroni familywise adjusted $p$-values (Grove and Andreasen 1982), where the $p$-value differs for the four families of variables (i.e., demographic characteristics, disordered eating, depression, and anxiety symptoms) according to the number of comparisons made. This resulted in an adjusted $p$-value for demographic characteristics of 0.017 (0.05/3), for disordered eating of 0.005 (0.05/10), for depression of 0.05 (0.05/1), and for anxiety of 0.025 (0.05/2). In the text, results are only discussed when group differences had $p$-values that reached significance using the appropriate Bonferroni corrected $p$-values.

Finally, we calculated Cohen's $d$ (Cohen 1988) effect sizes to determine the standardized magnitude of mean differences between twins and singletons on the dependent variables. An effect size of 0.20 is considered small, 0.50 moderate, and 0.80 large (Cohen 1988). Importantly, compensatory behaviors, eating concern, standardized depression scores, and BMI were log-transformed prior to analyses to adjust for positive skewness.

## Results

Descriptive information for twins and singletons is presented in Table 1. Twins were slightly older ($t = 10.96$, df = 1,284, $p < 0.001$) and had a slightly higher BMI ($t = 3.11$, df = 1,265, $p = 0.002$) than singletons. The mean (SD) age of the twins was 20.94 (2.46) years (range = 18.04–28.84 years), whereas the mean age of all the singletons was 19.27 (2.24) years (range = 17.00–51.00 years). In both twins and singletons, the majority of participants self-identified as being European–American (~83 %).

Though some statistically significant differences among the singleton samples emerged for unadjusted age, body dissatisfaction, restraint, and state anxiety values, only one reached statistical significance using a Bonferroni correction. Participants in singleton sample 3 were about 2 years older than the other samples ($p < 0.001$). Although the differences between samples were minimal, suggesting that the three singleton samples could be combined into one group, each singleton sample was separately compared to the twin sample in order to provide replication for testing the twin representativeness assumption.

Table 2 shows the unadjusted and adjusted means and standard errors for twins and each of the three singleton samples. For all three samples, twins had a lower mean score for all measures than singletons in both the unadjusted and adjusted models. We detected significant differences (adjusted $p$'s $\leq 0.002$) between both groups on most measures of disordered eating, with no significant differences for compensatory behaviors (adjusted $p$'s $\geq 0.12$) in any of the three singleton samples or body dissatisfaction in singleton sample 2 or singleton sample 3 (adjusted $p$'s $\geq 0.03$). In singleton sample 3, no significant mean differences emerged for the EDE-Q Global Score or the four subscales (adjusted $p$'s $\geq 0.01$). For depression, there was a significant mean difference between twins and singleton sample 1 (adjusted $p = 0.03$) but not between twins and singleton sample 2 (adjusted $p = 0.24$). For anxiety, there were significant mean differences between twins and singleton samples 1 and 2 for both state (adjusted $p$'s $< 0.001$) and trait (adjusted $p$'s $\geq 0.02$) anxiety.

We additionally examined whether there were significant variance differences between twin 1 and singletons. In general, the singletons exhibited more variability in their scores than the twins. This was true for the eating concern subscale in singleton samples 1 and 2 (all $p$'s $\leq 0.001$, with SD = 0.80 for twins and 1.14–1.29 for singletons), as well as EDE-Q restraint ($p$-value $\leq 0.003$, with SD = 1.33 for twins and 1.57 for singletons) and binge eating ($p = 0.001$, SD = 1.65 and 2.02 in twins and singletons, respectively) in singleton sample 1. This pattern of results followed those of the mean-level comparisons, such that singletons had more variability than twins, although differences appeared to be small. The only exception was for trait anxiety between twins and singleton sample 2 ($p < 0.001$), where twins reported slightly more variance than singletons on this measure (SD = 12.29 and 9.61 for twins and singletons, respectively).

Across all comparisons, the effect sizes for disordered eating, depression, and anxiety were between 0.02 and 0.38, which suggested that these mean differences were of

[2] Analyses were also conducted with twin 2 versus singletons, but the results were the same (results not shown). Therefore, we only reported results for twin 1.

**Table 1** Demographic characteristics

| | Twins ($n$ = 290–292) | Singleton sample 1 ($n$ = 468–474) | Singleton sample 2 ($n$ = 409–417) | Singleton sample 3 ($n$ = 103–105) |
|---|---|---|---|---|
| Age (Mean, SD) | 20.94 ± 2.46 | 18.90 ± 1.67 | 19.11 ± 2.09 | 21.61 ± 3.41 |
| Body mass index (Mean, SD) | 23.89 ± 5.35 | 22.86 ± 4.55 | 22.95 ± 3.85 | 23.35 ± 4.15 |
| Ethnicity (%) | | | | |
| European–American | 82.80 | 84.10 | 86.10 | 76.20 |
| African–American | 10.70 | 5.70 | 3.40 | 14.30 |
| Hispanic | 2.10 | 1.90 | 2.20 | 1.90 |
| Asian or Pacific Rim | 2.10 | 5.10 | 5.90 | 5.70 |
| Native American | 0.00 | 0.08 | 0.20 | 0.00 |
| Middle Eastern | 0.70 | 0.00 | 0.00 | 0.00 |
| Other | 1.70 | 2.30 | 2.20 | 1.90 |

*SD* standard deviation

small effect. The only exceptions were in the adjusted model for state anxiety with singleton sample 1 ($d = 0.51$) and binge eating with singleton sample 3 ($d = 0.47$), indicating a moderate effect size.

Finally, in case monozygotic and dizygotic twins systematically varied on these measures, we examined whether means and variances in disordered eating, depression, and anxiety scores differed by zygosity. As expected, there were no mean differences on any of these variables between monozygotic and dizygotic twins (all $p$-values $> 0.25$). In addition, similar to mean-level comparisons, there were no variance differences in disordered eating ($p$'s $\geq 0.03$), depression ($p > 0.05$), and anxiety symptoms ($p$'s $\geq 0.025$) between zygosity groups.

## Discussion

This study examined whether twins have similar rates of disordered eating, depression, and anxiety compared with singletons. In contrast to prior beliefs (Gau et al. 1992; Gjone and Novik 1995), twins reported less psychopathology than singletons across all three samples. Indeed, there were statistically significant differences in means and variances between twins and singletons for many of the disordered eating and anxiety measures, and for depression in one sample. However, these differences were of small effect and were consistent across all three singleton samples. Whether the mean and variance differences are clinically meaningful is unclear. In addition to effect sizes, the use of anchor-based methods, which compare quality of life assessments to measures that have clinical relevance (Crosby et al. 2003), will be an important future direction.

Our findings are consistent with prior research in suggesting that there are no significant mean differences for depression (Kendler et al. 1995, 1996; Pulkkinen et al. 2003), but possible mean and variance differences for disordered eating (Klump et al. 1999) and anxiety (Kendler et al. 1995), between twins and singletons. Still, the effect sizes were generally small, with twins reporting less psychopathology than singletons. Although Klump et al. (1999) and Kendler et al. (1995) found higher rates of disordered eating and anxiety in twins compared with singletons, certain differences between those studies and our study may account for the discrepant findings. First, in the Klump et al. (1999) study, the sample size was small (12 twins and 23 singletons). Our study was much larger ($n$ = 292 twins and 997 singletons). Second, we assessed state and trait anxiety rather than a specific measure of panic-phobia (Kendler et al. 1995). Although our measure of anxiety is likely related to panic-phobia, our results are not directly comparable to the Kendler et al. (1995) study given that different constructs of anxiety were assessed. More work is needed is understand whether the twin representativeness assumption is upheld for various anxiety constructs.

That twins had less psychopathology than singletons is consistent with "the adoptive hypothesis" (Pulkkinen et al. 2003), which suggests that being a twin actually has a positive impact on behavioral and psychological development. However, this hypothesis has not been directly examined to our knowledge. In addition, extended twin studies, which provide an alternative strategy to examine the twin representativeness assumption by including twins and their non-twin siblings, have not revealed significant effects of specific environmental influences shared by same-age siblings for disordered eating (Munn et al. 2010), depression, or generalized anxiety (Ehringer et al. 2006). The absence of this twin-specific shared environment suggests that any effects that are present are likely small and not clinically significant; thus, the results from twins may generalize to singleton populations. Taken together,

**Table 2** Unadjusted and adjusted means (standard error) for twins and singletons from hierarchical linear modeling

| | Unadjusted | | | | | Adjusted[b] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Twins | Singletons | *t* | *p* | Cohen's *d* | Twins | Singletons | *t* | *p* | Cohen's *d* |
| Singleton sample 1 | | | | | | | | | | |
| MEBS | | | | | | | | | | |
| Total score | 8.59 (0.46) | 11.29 (0.30) | −4.91 | **<0.001** | 0.38 | 8.40 (0.48) | 11.39 (0.31) | −4.91 | **<0.001** | 0.41 |
| Body dissatisfaction | 2.84 (0.16) | 3.38 (0.10) | −2.92 | **0.004** | 0.22 | 2.65 (0.16) | 3.47 (0.10) | −4.19 | **<0.001** | 0.35 |
| Compensatory behaviors[a] | 0.51 (0.07) | 0.66 (0.05) | −1.89 | 0.06 | 0.13 | 0.53 (0.07) | 0.66 (0.05) | −1.35 | 0.18 | 0.12 |
| Binge eating | 1.59 (0.13) | 2.32 (0.09) | −4.63 | **<0.001** | 0.35 | 1.59 (0.14) | 2.33 (0.09) | −4.13 | **<0.001** | 0.35 |
| Weight preoccupation | 3.03 (0.18) | 4.02 (0.12) | −4.75 | **<0.001** | 0.35 | 3.01 (0.19) | 4.02 (0.12) | −4.23 | **<0.001** | 0.36 |
| EDE-Q | | | | | | | | | | |
| Global score | 1.59 (0.10) | 1.97 (0.06) | −3.28 | **0.001** | 0.26 | 1.54 (0.10) | 1.99 (0.07) | −3.44 | **0.001** | 0.29 |
| Restraint | 1.28 (0.11) | 1.78 (0.07) | −4.00 | **<0.001** | 0.30 | 1.27 (0.12) | 1.78 (0.07) | −3.49 | **0.001** | 0.30 |
| Eating concern[a] | 0.61 (0.09) | 1.13 (0.06) | −5.08 | **<0.001** | 0.37 | 0.62 (0.09) | 1.13 (0.06) | −4.39 | **<0.001** | 0.37 |
| Shape concern | 2.13 (0.11) | 2.65 (0.07) | −3.78 | **<0.001** | 0.31 | 2.06 (0.12) | 2.67 (0.08) | −4.05 | **<0.001** | 0.33 |
| Weight concern | 1.78 (0.11) | 2.31 (0.07) | −3.94 | **<0.001** | 0.32 | 1.70 (0.12) | 2.35 (0.08) | −4.40 | **<0.001** | 0.35 |
| BDI | | | | | | | | | | |
| Total score[a] | −0.07 (0.08) | 0.09 (0.05) | −1.87 | 0.06 | 0.13 | −0.11 (0.08) | 0.12 (0.05) | −2.22 | **0.03** | 0.20 |
| STAI | | | | | | | | | | |
| State | 33.93 (0.77) | 40.23 (0.48) | −6.92 | **<0.001** | 0.55 | 33.91 (0.85) | 40.30 (0.52) | −6.04 | **<0.001** | 0.51 |
| Trait | 38.91 (0.85) | 41.41 (0.53) | −2.50 | **0.01** | 0.20 | 38.83 (0.92) | 41.50 (0.57) | −2.31 | **0.02** | 0.20 |
| Singleton sample 2 | | | | | | | | | | |
| MEBS | | | | | | | | | | |
| Total score | 8.59 (0.44) | 10.59 (0.31) | −3.71 | **<0.001** | 0.29 | 8.38 (0.44) | 10.81 (0.31) | −4.39 | **<0.001** | 0.36 |
| Body dissatisfaction | 2.84 (0.16) | 2.95 (0.11) | −0.59 | 0.55 | 0.04 | 2.68 (0.14) | 3.08 (0.10) | −2.25 | 0.03 | 0.19 |
| Compensatory behaviors[a] | 0.51 (0.06) | 0.65 (0.05) | −1.70 | 0.09 | 0.14 | 0.52 (0.07) | 0.65 (0.05) | −1.58 | 0.12 | 0.12 |
| Binge eating | 1.59 (0.12) | 2.21 (0.09) | −4.05 | **<0.001** | 0.32 | 1.58 (0.13) | 2.24 (0.09) | −4.05 | **<0.001** | 0.34 |
| Weight preoccupation | 3.03 (0.17) | 3.89 (0.12) | −4.12 | **<0.001** | 0.33 | 2.97 (0.18) | 3.93 (0.13) | −4.27 | **<0.001** | 0.35 |
| EDE-Q | | | | | | | | | | |
| Global score | 1.59 (0.09) | 2.04 (0.07) | −3.97 | **<0.001** | 0.31 | 1.53 (0.10) | 2.09 (0.07) | −4.73 | **<0.001** | 0.37 |
| Restraint | 1.28 (0.10) | 1.66 (0.07) | −3.10 | **0.002** | 0.25 | 1.26 (0.11) | 1.68 (0.07) | −3.10 | **0.002** | 0.26 |
| Eating concern[a] | 0.61 (0.08) | 0.96 (0.05) | −3.71 | **<0.001** | 0.30 | 0.60 (0.08) | 0.97 (0.06) | −3.80 | **<0.001** | 0.29 |
| Shape concern | 2.13 (0.12) | 2.69 (0.08) | −4.00 | **<0.001** | 0.31 | 2.03 (0.11) | 2.76 (0.08) | −5.08 | **<0.001** | 0.43 |
| Weight concern | 1.78 (0.11) | 2.25 (0.08) | −3.52 | **<0.001** | 0.27 | 1.70 (0.11) | 2.30 (0.07) | −4.47 | **<0.001** | 0.38 |
| BDI | | | | | | | | | | |
| Total score[a] | −0.07 (0.07) | 0.00 (0.05) | −0.84 | 0.40 | 0.06 | −0.09 (0.08) | 0.02 (0.05) | −1.17 | 0.24 | 0.10 |
| STAI | | | | | | | | | | |
| State | 33.93 (0.73) | 38.76 (0.49) | −5.47 | **<0.001** | 0.44 | 33.98 (0.77) | 38.74 (0.52) | −4.96 | **<0.001** | 0.41 |
| Trait | 38.91 (0.75) | 42.02 (0.51) | −3.44 | **0.001** | 0.27 | 38.76 (0.79) | 42.22 (0.54) | −3.53 | **<0.001** | 0.29 |
| Singleton sample 3 | | | | | | | | | | |
| MEBS | | | | | | | | | | |
| Total score | 8.59 (0.44) | 10.38 (0.61) | −2.39 | 0.02 | 0.25 | 8.48 (0.42) | 10.85 (0.60) | −3.23 | **0.001** | 0.35 |
| Body dissatisfaction | 2.84 (0.16) | 2.90 (0.22) | −0.22 | 0.83 | 0.02 | 2.78 (0.14) | 3.04 (0.20) | −1.09 | 0.28 | 0.12 |
| Compensatory behaviors[a] | 0.51 (0.06) | 0.57 (0.09) | −0.51 | 0.61 | 0.06 | 0.50 (0.06) | 0.62 (0.10) | −0.98 | 0.33 | 0.12 |
| Binge eating | 1.59 (0.12) | 2.38 (0.18) | −3.64 | **<0.001** | 0.40 | 1.57 (0.12) | 2.50 (0.18) | −4.20 | **<0.001** | 0.47 |
| Weight preoccupation | 3.03 (0.17) | 3.89 (0.24) | −3.00 | **0.003** | 0.31 | 2.98 (0.16) | 3.98 (0.24) | −3.43 | **0.001** | 0.38 |
| EDE-Q | | | | | | | | | | |
| Global score | 1.59 (0.09) | 1.80 (0.12) | −1.37 | 0.17 | 0.14 | 1.55 (0.09) | 1.89 (0.12) | −2.36 | 0.02 | 0.24 |
| Restraint | 1.28 (0.09) | 1.37 (0.13) | −0.59 | 0.56 | 0.06 | 1.25 (0.09) | 1.40 (0.13) | −0.89 | 0.38 | 0.10 |
| Eating concern[a] | 0.61 (0.07) | 0.86 (0.10) | −2.07 | 0.04 | 0.22 | 0.60 (0.07) | 0.92 (0.10) | −2.75 | 0.01 | 0.28 |

**Table 2** continued

| | Unadjusted | | | | | Adjusted[b] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Twins | Singletons | *t* | *p* | Cohen's *d* | Twins | Singletons | *t* | *p* | Cohen's *d* |
| Shape concern | 2.13 (0.11) | 2.37 (0.15) | −1.23 | 0.22 | 0.13 | 2.07 (0.11) | 2.49 (0.15) | −2.33 | 0.02 | 0.24 |
| Weight concern | 1.78 (0.11) | 2.03 (0.15) | −1.39 | 0.17 | 0.14 | 1.73 (0.10) | 2.16 (0.14) | −2.56 | 0.01 | 0.27 |

*p*-values that remained significant after Bonferroni familywise correction (disordered eating symptoms: $p < 0.005$; depression symptoms: $p < 0.05$; anxiety symptoms: $p < 0.025$) are in bold-type

*MEBS* Minnesota Eating Behavior Survey, *EDE-Q* Eating Disorder Examination Questionnaire, *BDI* Beck Depression Inventory, *STAI* State-Trait Anxiety Inventory

[a] Raw/standardized means and standard errors are reported, with significance reported according to log-transformed scores

[b] Adjusted for age, body mass index, and ethnicity

findings from this study and extended twin studies that include twins and their non-twin siblings suggest that twins and singletons do not differ substantially on levels of disordered eating and internalizing symptoms. It is even possible that being a twin has a positive impact on reducing these symptoms.

Findings from the current and prior studies are important as they broaden our ability to integrate research on genetic risk factors for these behaviors from both twin-based and population-based samples. This conclusion is especially important when discussing the heritability of a behavior. The most popular method for determining heritability is via the classical twin design, in which the correlation between monozygotic same-sex twins is compared with the correlation between dizygotic same-sex twins to determine the magnitude of genetic and environmental effects on the phenotype. These methods cannot be applied solely to non-twin individuals. Traditional epidemiological methods suggest that before searching for genetic markers that are associated with a behavior, it is important to establish that the behavior is heritable. Although family-based association studies (which can include twin pairs) are informative to search for genetic variants, there are many criticisms for using this approach as opposed to using population-based (i.e., singleton) samples, including the increased genotyping cost and difficulty in recruiting entire families as opposed to one individual from a family (Hopper et al. 2005). Still, there are many long-standing twin registries containing a wealth of genotypic and environmental data that will be increasingly important in elucidating the genetic and environmental influences on psychopathology (van Dongen et al. 2012), as they can provide additional information that singleton samples cannot, such as having a matched control for environmental factors shared by family members (Hopper et al. 2005).

There are a number of strengths associated with this study. First, it is one of the largest studies to date to examine the twin representativeness assumption for eating pathology. Second, a sophisticated statistical approach was used that enabled us to compare twin and singleton self-reports while accounting for the non-independence of twin data. Third, findings were consistent across three different singleton samples. Lastly, all participants were recruited from the same geographic area, which controlled for potential regional biases associated with disordered eating, depression, and anxiety.

Despite these strengths, there are some limitations worth noting. First, participants self-reported symptoms and did not receive clinical diagnoses of eating disorders, depression, or anxiety. Therefore, results may not generalize to clinical populations. Second, all individuals may not have completed the period of risk for eating pathology, depression, and anxiety, which can begin in early adulthood (American Psychiatric Association 2000). For example, research has shown that some disordered eating symptoms (i.e., binge eating and compensatory behaviors) typically onset before age 19 (Stice et al. 1998); this is also the mean age of the twins and singletons in this sample. Thus, although some of the participants may not yet have traversed the period of risk for clinical eating disorders, depression, or anxiety disorders, the variables we examined represent risk factors for clinical disorders (Eaton et al. 2013; Hankin et al. 2005; Striegel-Moore et al. 1986). Therefore, our findings likely generalize to these disorders and their risk trajectories. Third, height and weight were measured differently across the groups: in twins, it was directly measured by trained research assistants and in singletons, it was obtained through self-report. The difference in data collection methods could have introduced some systematic bias such that the BMIs of singletons were reported to be lower than they were. Prior research has shown that female adolescents and young adults tend to under-report their weight by an average of 2.1 pounds; however, this is likely due to random error (Field et al. 2007). Still, the difference in BMI reporting methods did not seem to have impacted our findings, as the pattern of results for both the unadjusted and adjusted means

for disordered eating were similar to those for depression and anxiety (Table 2). Fourth, the findings are sample- and measure- specific. Additional studies that use other disordered eating, depression, and anxiety assessments are needed. Finally, our singletons were not recruited from the general population, but rather through a university. Indeed, nearly 50 % of the twins were recruited in a similar manner. Future research should examine differences in psychopathology between twins and singletons recruited from the larger population.

In conclusion, although there were no significant mean differences between twins and singletons in compensatory behaviors (and depression in one comparison), this study found statistically significant differences between both groups for most measures of disordered eating and anxiety, with twins reporting less psychopathology than singletons. It is unknown whether these differences are clinically meaningful. Future studies are needed to determine whether the twin representativeness assumption holds for measures of disordered eating and internalizing symptoms, and the degree to which any differences hold clinical significance. In addition, future work should examine possible sex, age, education level, and ethnic/racial differences for this assumption, as the findings reported in this study may not generalize to men, adolescents, older adults, or non-Caucasian ethnicities.

**Acknowledgments** The authors are grateful for the assistance of Jessica L. Suisman, M.A. and Yunfei Wu, M.S. This work was funded by T32 AA07580 to MMC, as well as T32 MH070343 and T32 MH082761 to KMC.

## References

American Psychiatric Association (2000) Diagnostic and statistical manual of mental disorders, 4th edn, text revision. American Psychiatric Press, Washington, DC

Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J (1961) An inventory for measuring depression. Arch Gen Psychiatry 4:561–571

Beck AT, Steer RA, Garbin MG (1988) Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. Clin Psychol Rev 8:77–100

Beck AT, Steer RA, Brown GK (1996) Beck Depression Inventory Manual, 2nd edn. Psychological Corporation, San Antonio

Burt SA, Klump KL (2013) The Michigan State University Twin Registry (MSUTR): an update. Twin Res Hum Genet 16: 344–350

Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Lawrence Earlbaum Associates, Hillsdale

Crosby RD, Kolotkin RL, Williams GR (2003) Defining clinically meaningful change in health-related quality of life. J Clin Epidemiol 56:395–407

Culbert KM, Breedlove SM, Burt SA, Klump KL (2008) Prenatal hormone exposure and risk for eating disorders: a comparison of opposite-sex and same-sex twins. Arch Gen Psychiatry 65: 329–336

Dozois DJ, Dobson KS, Ahnberg JL (1998) A psychometric evaluation of the Beck Depression Inventory-II. Psychol Assess 10:83–89

Eaton NR, Krueger RF, Markon KE, Keyes KM, Skodol AE, Wall M, Hasin DS, Grant BF (2013) The structure and predictive validity of the internalizing disorders. J Abnorm Psychol 122:86–92

Ehringer MA, Rhee SH, Young S, Corley R, Hewitt JK (2006) Genetic and environmental contributions to common psychopathologies of childhood and adolescence: a study of twins and their siblings. J Abnorm Child Psychol 34:1–17

Evans DM, Martin NG (2000) The validity of twin studies. GeneScreen 1:77–79

Fairburn CG, Beglin SJ (1994) Assessment of eating disorders: interview or self-report questionnaire? Int J Eat Disord 16: 363–370

Field AE, Aneja P, Rosner B (2007) The validity of self-reported weight change among adolescents and young adults. Obesity 15:2357–2364

Gau JS, Silberg JL, Erickson MT, Hewitt JK (1992) Childhood behavior problems: a comparison of twin and non-twin samples. Acta Genet Med Gemellol 41:53–63

Gjone H, Novik TS (1995) Parental ratings of behaviour problems: a twin and general population comparison. J Child Psychol Psychiatry 36:1213–1224

Grove WM, Andreasen NC (1982) Simultaneous tests of many hypotheses in exploratory research. J Nerv Ment Dis 170:3–8

Hankin BL, Fraley RC, Lahey BB, Waldman ID (2005) Is depression best viewed as a continuum or discrete category? A taxometric analysis of childhood and adolescent depression in a population-based sample. J Abnorm Psychol 114:96–110

Hopper JL, Bishop DT, Easton DF (2005) Population-based family studies in genetic epidemiology. Lancet 366:1397–1406

Iacono WG, Carlson SR, Taylor J, Elkins IJ, McGue M (1999) Behavioral disinhibition and the development of substance-use disorders: findings from the Minnesota Twin Family Study. Dev Psychopathol 11:869–900

IBM (2011) SPSS version 20. IBM, Armonk

Kendler KS (1993) Twin studies of psychiatric illness. Current status and future directions. Arch Gen Psychiatry 50:905–915

Kendler KS, Martin NG, Heath AC, Eaves LJ (1995) Self-report psychiatric symptoms in twins and their nontwin relatives: are twins different? Am J Med Genet 60:588–591

Kendler KS, Pedersen NL, Farahmand BY, Persson PG (1996) The treated incidence of psychotic and affective illness in twins compared with population expectation: a study in the Swedish Twin and Psychiatric Registries. Psychol Med 26:1135–1144

Kendler KS, Gatz M, Gardner CO, Pedersen NL (2006) A Swedish national twin study of lifetime major depression. Am J Psychiatry 163:109–114

Klump KL, Burt SA (2006) The Michigan State University Twin Registry (MSUTR): genetic, environmental and neurobiological influences on behavior across development. Twin Res Hum Genet 9:971–977

Klump KL, Keel PK, Leon GR, Fulkerson JA (1999) Risk for eating disorders in a school-based twin sample: are twins representative of the general population for eating disordered behavior? Eat Disord 7:33–41

Klump KL, McGue M, Iacono WG (2000) Age differences in genetic and environmental influences on eating attitudes and behaviors in preadolescent and adolescent female twins. J Abnorm Psychol 109:239–251

Lykken DT, Bouchard TJ Jr, McGue M, Tellegen A (1990) The Minnesota Twin Family Registry: some initial findings. Acta Genet Med Gemellol 39:35–70

Mond JM, Hay PJ, Rodgers B, Owen C, Beumont PJ (2004) Temporal stability of the Eating Disorder Examination Questionnaire. Int J Eat Disord 36:195–203

Mosing MA, Gordon SD, Medland SE, Statham DJ, Nelson EC, Heath AC, Martin NG, Wray NR (2009) Genetic and environmental influences on the co-morbidity between depression, panic disorder, agoraphobia, and social phobia: a twin study. Depress Anxiety 26:1004–1011

Munn MA, Stavro GM, Klump KL (2005) Lack of seasonal variation in eating attitudes and behaviours among female college students. Eur Eat Disord Rev 13:101–105

Munn MA, Stallings MC, Rhee SH, Sobik LE, Corley RP, Rhea SA, Hewitt JK (2010) Bivariate analysis of disordered eating characteristics in adolescence and young adulthood. Int J Eat Disord 43:751–761

Peeters H, Van GS, Vlietinck R, Derom C, Derom R (1998) Validation of a telephone zygosity questionnaire in twins of known zygosity. Behav Genet 28:159–163

Peterson CB, Crosby RD, Wonderlich SA, Joiner T, Crow SJ, Mitchell JE, Bardone-Cone AM, Klein M, Le Grange D (2007) Psychometric properties of the Eating Disorder Examination-Questionnaire: factor structure and internal consistency. Int J Eat Disord 40:386–389

Pulkkinen L, Vaalamo I, Hietala R, Kaprio J, Rose RJ (2003) Peer reports of adaptive behavior in twins and singletons: is twinship a risk or an advantage? Twin Res 6:106–118

Racine SE, Culbert KM, Larson CL, Klump KL (2009) The possible influence of impulsivity and dietary restraint on associations between serotonin genes and binge eating. J Psychiatr Res 43:1278–1286

Spielberger CD, Gorsuch RL, Lushene RE, Vagg RE, Jacobs GA (1983) Manual for the state-trait anxiety inventory. Consulting Psychologists Press, Palo Alto

Stice E, Killen JD, Hayward C, Taylor CB (1998) Age of onset for binge eating and purging during late adolescence: a 4-year survival analysis. J Abnorm Psychol 107:671–675

Striegel-Moore RH, Silberstein LR, Rodin J (1986) Toward an understanding of risk factors for bulimia. Am Psychol 41:246–263

Thornton LM, Mazzeo SE, Bulik CM (2011) The heritability of eating disorders: methods and current findings. Curr Top Behav Neurosci 6:141–156

van Dongen J, Slagboom PE, Draisma HH, Martin NG, Boomsma DI (2012) The continuing value of twin studies in the omics era. Nat Rev Genet 13:640–653

von Ranson KM, Klump KL, Iacono WG, McGue M (2005) The Minnesota Eating Behavior Survey: a brief measure of disordered eating attitudes and behaviors. Eat Behav 6:373–392