Behavior Genetics, Vol. 30, No. I, 2000

Physical Similarity and Twin Resemblance for Eating Attitudes and Behaviors: A Test of the Equal Environments Assumption

Kelly L. Klump,[1,3] Amanda Holly,[2] William G. Iacono,[2] Matt McGue,[2] and Laura E. Willson[2]

Received 3 Nov. 1998—Final 20 Sept. 1999

The Equal Environments Assumption (EEA) in twin studies of eating pathology was investigated by examining the hypothesis that twin resemblance for eating attitudes and behaviors is affected by their degree of physical similarity. Eating attitudes and behaviors were assessed in 338 female adolescent twin pairs with a revised version of the Eating Disorder Inventory (EDI). General physical similarity as well as body size/shape similarity were assessed using ratings of color photographs, ratings of body shape, and body mass index. All physical similarity assessments were conducted blind to twin zygosity. Significant associations between physical similarity and twin similarity for eating attitudes and behaviors were not found. Mean EDI within-twin pair absolute difference scores did not differ significantly among more versus less physically-similar groups. Additionally, correlation and regression analyses failed to find a significant association between ED1 absolute difference scores and physical similarity indices. The current findings provide support for the EEA in twin studies of eating attitudes and behaviors.

KEY WORDS: Equal environments assumption; physical similarity; twin study; eating pathology; adolescent.

INTRODUCTION

One of the most popular behavioral genetic designs for examining genetic and environmental influences on phenotypes has been the classical twin study. The usefulness of this design depends upon the validity of multiple assumptions on which it is based. One of the assumptions, the Equal Environments Assumption (EEA), has been continually challenged by twin study critics. The EEA. posits that both members of a monozygotic (MZ) twin pair are as likely to be treated the same by the environment as both members of a dizygotic (DZ) twin pair with respect to environmental influences that are of etiologic importance to the trait under study (Plomin *et al.,* 1990). If MZ and DZ twins are not equally correlated for important environmental factors, than greater MZ relative to DZ twin resemblance for disorders may be due to environmental in addition to genetic variance.

Twin study critics have argued for the invalidity of the EEA based on three related beliefs: (1) MZ twins are physically more similar than DZ twins; (2) MZ twins are treated more similarly than DZ twins; and (3) greater environmental similarity leads to increased MZ relative to DZ twin resemblance. Behavioral genetic research has confirmed the first two of these perceptions: MZ twins have been shown to be physically more similar than DZ twins, and they have been found to experience greater environmental similarity (see Plomin et *al.,* 1990, for a review). However, for most traits and disorders studied, the third belief has been

[1] Department of Psychology, Michigan State University, East Cansing, Michigan.
[2] Department of Psychology, University of Minnesota, Minneapolis, Minnesota.
[3] To whom correspondence should be addressed at Department of Psychology, Michigan State University, 129 Psychology Research Building, East Cansing, MI 48824. Fax (517) 432-2476.

0001-8244/00/0100-005 IS 18.00/0

disconfirmed by research showing that greater MZ twin environmental similarity has not been directly related to increased MZ relative to DZ twin resemblance for personality (Loehlin and Nichols, 1976; Matheny *et al.*, 1976; Plomin, 1976; Scarr and Carter-Saltzman, 1979), cognitive abilities (Matheny *et al.*, 1976; Scarr and Carter-Saltzman, 1979), or psychiatric disorders (Hettema *et al.*, 1995; Kendler, 1983; Kendler *et al.*, 1993; Kendler *et al.*, 1992a; Kendler *et al.*, 1992b; Kendler *et al.*, 1992c; Kendler *et al.*, 1992d). For example, Loehlin and Nichols (1976) found that MZ twins were rated significantly more alike than DZ twins on six indices of differential experience in childhood, including the extent to which the twins played together, were dressed alike, spent time together, slept in the same room, were treated similarly by parents, and had the same teachers in school. However, although the MZ twins were more similar on all of these indices, Loehlin and Nichols (1976) did not find that these more similar environments were associated with greater personality similarity in MZ relative to DZ twins. Correlating rated similarity with within-pair difference scores on 18 scales of the California Personality Inventory, the authors found that MZ twins who were treated more similarly by their environment were no more similar in personality than MZ twins who were treated less similarly. The authors concluded that the assumption that greater similarity in environmental treatment accounts for later personality similarity in MZ twins was not warranted.

One psychiatric disorder that may provide an exception to the generalizability of the EEA is bulimia nervosa (BN). A series of studies using the same epidemiological sample has provided both positive and negative support for the EEA. The validity of the EEA for bulimic twins has been upheld in these studies in terms of perceived versus real zygosity (Kendler *et al.*, 1993), similarity of childhood treatment (Bulik *et al.*, 1998; Kendler *et al.*, 1991; Sullivan *et al.*, 1998), and frequency of contact as adults (Kendler *et al.*, 1991). However, two studies by this research group obtained discrepant results with regard to the effect of physical similarity on the validity of the EEA. These investigations examined the hypothesis that twin resemblance for BN is affected by the similarity of their environmental treatment, which in turn is a result of their degree of physical similarity. Using twin model-fitting analyses and a categorical assessment of BN (i.e., affected/unaffected), Hettema *et al.* (1995) found general physical similarity, assessed from waist-up color photographs, to account for 64–72% of the variance in BN. This finding was unique to BN; the authors did not find physical similarity to account for significant proportions of variance in other disorders examined including major depression, generalized anxiety disorder, phobia, and alcoholism.

Hettema *et al.* (1995) concluded that because BN has been linked to "slimmer ideal body image, low levels of self-esteem, and an external locus of control" it therefore "may not be surprising that twins with a greater degree of physical resemblance, and therefore similar levels of body image dissatisfaction and appearance-based self-esteem, have a greater risk for bulimia or bulimia-type behavior than those who are less similar in appearance" (p. 333). However, in later studies of the heritability of BN and behaviors comprising BN (i.e., bingeing and purging) using the same subject sample and the same photograph ratings, Bulik et *al.*, (1998) and Sullivan et *al.* (1998) did not find a relationship between physical similarity and twin similarity for these behaviors. Controlling for zygosity, logistical regression analyses failed to find an association between physical similarity and twin concordance/discordance for BN, bingeing, or purging.

Discrepant findings between studies may be due to a number of methodological limitations. First, there was a relatively high correlation (r = 0.78) between zygosity and physical similarity ratings, due in part to the fact that the algorithm used to diagnose zygosity used the same photographs as those rated for physical similarity. This may have limited the authors' ability to separate the effects of these two variables. Second, interrater reliability for the photograph ratings was modest (r = 0.68), potentially leading to inconsistent results. Third, the relatively small number of bulimic twins (n = 99) in the studies may have influenced findings, especially given the large sample sizes needed for twin model-fitting analyses. With smaller sample sizes, the standard error associated with these parameter estimates increases (Neale and Cardon, 1992).

A final potential limitation of Hettema *et al. (1995)* is the confounding of physical similarity with clinical state in bulimic twins. That is, twins who are concordant for BN may be more physically similar than discordant twins because of the effects of the disorder on body weight and shape in both twins. The use of cross-sectional data in these studies makes it difficult to determine whether appearance similarity was present before the onset of the disorder, and thus contributed to twin similarity for BN. Employing longitudinal designs or examining the influence of physical similarity on bulimic attitudes and behaviors in nonclinical samples may aid in this differentiation.

The purpose of the present investigation was to examine the hypothesis that twin resemblance for BN is affected by the similarity of their environmental treatment, which in turn is a result of their degree of physical similarity. Methodological improvements over previous research were incorporated whenever possible. For example, continuous measures of eating pathology were used to increase statistical power. In addition, these measures were administered to a large, population-based sample of twins to avoid confounding clinical state with physical similarity. Further, a more comprehensive assessment of physical similarity was used. Measures of both general physical similarity as well as body size resemblance were used in an attempt to identify important aspects of physical similarity. For eating pathology, body size similarity may exert a greater influence on twin concordance than general resemblance. Finally, physical similarity measurements that are relatively independent of zygosity diagnosis and classification were employed.

METHOD

Participants

The sample used in the present study comprised female twin participants in the Minnesota Twin Family Study (MTFS). The MTFS is a population-based, longitudinal study of the development of substance use and related disorders in reared-together male and female twins and their parents (Iacono et *al.*, 1996; McGue et *al.*, 1996). Twin families were ascertained from State of Minnesota birth records and located using various public databases including telephone directories and driver license registrations (Lykken et *al.*, 1990). For any given birth year, the MTFS has been able to locate more than 90% of twin births. Families were excluded from participation if they lived further than a day's drive from our laboratories in Minneapolis, if one member of the twin pair was deceased, if the twins were adopted, or if either of the twins in the family had a serious physical or cognitive handicap that would preclude him/her from completing the day-long, in-person intake assessment. Among those families who were eligible and were asked to participate, approximately one-sixth refused. Because both MTFS participants and most nonparticipants completed a mailed questionnaire covering family demographics and health, it was possible to compare these groups on these variables. The results of these comparisons, which have been summarized elsewhere (Iacono et *al.*, 1999), revealed only minor differences between these groups (e.g., participating families had slightly higher socioeconomic status than nonparticipants).

The initial subject sample included 338 sixteen, seventeen, and eighteen year-old female twin pairs. Fifty two (15.4%) of the twin pairs were excluded from analyses due to one or both co-twins of a twin Pair having missing data on either the physical similarity indices or the measure of eating attitudes and behaviors, An entire twin pair was excluded if only one twin was missing data because data for both twins must be present in order to assess within-twin pair similarity. The most common reasons for missing data were computer malfunctions, leading to an inability to administer the computer-based assessment of eating attitudes and behaviors, and an inability to assess general physical similarity due to missing photographs or poor quality of existing photographs. Less common reasons for missing data include: (1) the failure of twins to return body shape ratings by mail when they were unable to finish them during the day-long, in-person assessment; and (2) twins' refusal to be weighed and/or to have their height measured, leading to an inability to assess relative body weight. Notably, none of the twins with missing data who had completed the MTFS's structured diagnostic interview for DSMIV diagnoses of anorexia nervosa and BN met criteria for either of these eating disorders, suggesting that the exclusion of subjects with missing data did not significantly influence results. The final twin sample comprised 286 (84.6%) female twin pairs (mean age = 17.46 years, SD = 0.51).

Zygosity Determination

Zygosity was established using: (1) a parental zygosity questionnaire regarding the physical similarity of the twins; (2) a project zygosity estimate consisting of a trained research assistant's face-to-face evaluation of the twins' similarity in eye color and ear configuration; and (3) an algorithm diagnosis calculated from ponderal index, cephalic index, and fingerprint ridge count. The rate of disagreement among these three indices in the entire 17 year-old twin sample was 33.7% (117/338 twin pairs). These disagreements were resolved through a serological examination of 12 blood group antigens and protein polymorphisms. In a validation study using 50 pairs of twins all of whom had the serological test, in every case where the three zygosity estimates agreed, the serological analysis confirmed the agreement.

Measures

Eating Attitudes and Behaviors. Self-reported eating attitudes and behaviors were assessed with a computer-administered, 30-item revised Eating Disorders Inventory (EDI; Garner *et al.,* 1983). A previous factor analysis (Klump *et al.,* 1999) of the revised EDI yielded the following four subscales: Body Dissatisfaction (dissatisfaction with the size and/or shape of one's body), Weight Preoccupation (preoccupation with dieting, weight, and the pursuit of thinness), Binge Eating (the tendency to engage in episodes of overeating as well as having attitudes conducive to binge eating), and Compensatory Behavior (the tendency to use or to contemplate using inappropriate compensatory behaviors such as self-induced vomiting and laxatives to control weight). The ED1 and its subscales are scored in the traditional "pathological" direction with high scores indicating greater degrees of the measured construct. A previous study of the psychometric properties of the revised ED1 has supported its reliability and validity (Klump *et al.,* 1998).

ED1 Total Scores were prorated for 4 (0.6%) twins who were missing one to three (<10%) of the ED1 items. Given the relatively small number of items on each subscale (range = 6–8 items), data from participants with one or two missing items on any given subscale were not included in analyses of that scale. Differing sample sizes for subscale analyses reflect this exclusion.

Physical Similarity. As mentioned above, physical similarity was measured two ways in order to increase the comprehensiveness of the assessment. First, body weight and shape were assessed with a "Physical Size Index" (PSI). The PSI is a summed combination of the absolute value of standardized, within-pair difference scores for body mass index (BMI = weight (kg)/height^2 (m)) plus the absolute value of standardized, within-pair difference scores for body shape ratings. The high and significant correlation ($r = .83$; $p < .001$) between BMI and body shape ratings indicated the appropriateness of combining them to form a single index.

For BMI estimates, height was measured with an anthropometer and weight was measured on a level platform scale with a beam and moveable weights. Ratings of body shape were made with the MTFS Body Rating Scales (BRS; Sherman *et al.,* 1995). The BRS are line drawings of nine body silhouettes ranging from thin (1) to obese (9). These scales were filled out by clinical interviewers who were unaware of the twins' measured height and weight at the time of rating. A previous study of the psychometric properties of the MTFS BRS scales has supported their reliability and validity (Sherman *et al.,* 1995).

Second, ratings of general physical similarity were made using photographs of each twin. Multiple photographs of twins and their families were taken during their day-long intake assessment at the MTFS. Standard protocol dictated that at least one waist-up and one full-length photo of each twin, one picture of the twins together, and one picture of the twins with their parents be taken during the visit. Waist-up photographs of each twin were chosen for rating in the current study to decrease bias resulting from similar posturing and gesturing when the twins had their picture taken together. However, when the waist-up photograph was not available, one of the other types of photographs was rated in its place (n = 45 (15.7%) twin pairs).

Two trained female undergraduate research assistants independently rated the twin photographs on a scale of 1 (least similar; no familial resemblance) to 5 (most similar; twins look like "two peas in a pod"). Twin differences in hair color and style, hygiene, and other environmentally-manipulated characteristics were taken into account, as such differences are likely to lead to perceived differences in general appearance. Clothing differences were not taken into account, however, as these differences may reflect day-to-day changes. These rating standards were similar to those used in the Hettema *et al.* (1995) study (K. S. Kendler, personal communication, January, 1997). All ratings were conducted blind to the MTFS zygosity determination.

Following the initial rating of photographs, the two research assistants met to discuss cases in which their ratings differed. A "consensus" rating was then assigned to each discrepant case.

Interrater reliability, test-retest reliability, and stability statistics for the pre-consensus photograph ratings were calculated with intraclass correlations. Two-week test-retest reliability was calculated for 95 pairs as the intraclass correlation between a rater's initial ratings on twin photographs and the rater's score on the same photographs two weeks later. A three-year stability statistic was also calculated for 31 twin pairs as the intraclass correlation between a rater's initial ratings on the twin photograph and the rater's score for the same twin pair using different photographs that had been taken when the twins were approximately three years older. This statistic provides information regarding both the stability of rated similarity as well as the nature (state versus trait) of the measured characteristics.

Interrater reliability (r = .77), two week test-retest reliability (r = .96), and the stability statistic (r = .70) were all moderate-to-high. These findings suggest that general physical similarity can be reliably scored from photographs by independent raters, that the photograph ratings are reliable and stable over time, and that the ratings tap trait rather than state characteristics of physical similarity.

As indicated above, photographs other than waist-up photographs of each twin were used, for 45 twin pairs for whom the standard photograph was unavailable. Consequently, in order to assess comparability of ratings when different photographs were used, intraclass correlations were calculated for 39 twin pairs for whom both the standard photograph as well as a singular photograph of the twins together were rated. Two types of reliability statistics were calculated: (1) interrater reliability for the nonstandard photograph, or the correlation between Rater l's and Rater 2's rating on the nonstandard, twins together photograph; and (2) within-rater reliability across photographs, or the correlation between each rater's scores on the standard photograph with each rater's scores on the nonstandard, twins together photograph. The moderate-to-high correlations for both the interrater reliability (r = .71) as well as the within-rater reliabilities (Rarer l's r = .86; Rater 2's r = .90) indicate considerable reliability and consistency in ratings across photographs. These findings suggest that ratings of different photographs are comparable and that the inclusion of the substitute photographs should not have had any appreciable effect on study results.

Statistical Analyses

All statistical analyses were conducted within zygosity. MZ and DZ twins were analyzed separately because observed differences among MZ twins are clearly environmental in origin, while observed differences among DZ twins may be genetic or environmental. Keeping the two types of twins separate allows for clearer interpretation of data.

The relationship between physical similarity and twin resemblance for EDI scores was examined in two ways. First, twin pairs were stratified according to their photograph rating. Mean absolute within-pair difference scores for EDI scales were then compared across photograph ratings using univariate analysis of variance (ANOVA) and a one-factor multivariate analysis of variance (MANOVA). The MANOVA factor was photograph ratings (ratings 1–5), with mean within-pair difference scores for the ED1 *subscales* serving as dependent variables. As the EDI Total Score is a linear composite of its subscales, it was not included in the MANOVA, but was instead included in a separate ANOVA, with mean within-pair difference scores for EDI Total Score as the dependent variable and photograph ratings (ratings 1–5) as independent variables. Significantly smaller mean differences within physically-similar relative to physically-dissimilar twin pairs for ED1 scales would suggest an effect of general physical resemblance on twin similarity for eating pathology.

Second, Pearson product moment correlations were calculated between the PSI and ED1 within-pair absolute difference scores (Plomin et al., 1976). Significant positive correlations would indicate an association between body weight and shape similarity and twin resemblance for eating attitudes and behaviors.

RESULTS

Table I presents the means and standard deviations for the EDI absolute difference scores and physical similarity indices. The EDI and BMI data were transformed ($\log_{10}(x + 1)$) prior to analysis due to the positive skew of the data. Difference scores for these measures were calculated using the log-transformed values. For descriptive purposes, raw EDI scores, calculated from the non-normal data, are reported in tables.

Three conclusions may be drawn from these data. First, MZ twins showed greater resemblance for eating attitudes and behaviors compared to DZ twins, a result that is similar to previous findings in this area (Holland et al., 1988; Klump et *al.*, 2000; Rutherford et *al.*, 1993; Wade et *al.*, 1998). Second, as expected, MZ twins were rated as significantly more physically similar than DZ twins on both measures. Finally, there is variation in similarity of physical appearance in both MZ and DZ twins. This last point is crucial for the present study. If there were no variability in appearance among MZ and DZ twins, then examining the extent to which physical resemblance relates to similarity in eating pathology would be meaningless.

Tables II and III present results from analyses examining the relationship between physical similarity and twin resemblance for EDI scores. Overall, no clear significant association between physical appearance and twin similarity for eating attitudes and behaviors was found. For example, Table II presents the mean within-pair absolute difference scores for EDI scales, stratified according to photograph ratings. A nonsignificant ANOVA in both MZ ($F(2,186) = 1.04, p = .35$) and

Table I. Means and (Standard Deviations) for EDI Absolute Difference Scores and Physical Similarity Indices[a]

Measures	Total sample (n = 284–286)	MZ (n = 185–186)	DZ (n = 98–100)
EDI difference scores			
EDI total score	4.58 (3.77)	3.91 (3.03)	5.87 (4.60)**
Body dissatisfaction	1.65 (1.62)	1.34 (1.39)	2.25 (1.85)**
Compensatory behavior	0.03 (1.04)	0.03 (0.94)	0.10 (1.18)
Binge eating	1.35 (1.29)	1.20 (1.25)	1.61 (1.30)*
Weight preoccupation	2.03 (1.79)	1.69 (1.57)	2.63 (2.00)**
Physical similarity indices			
Photograph ratings	3.80 (1.08)	4.34 (0.64)	2.75 (0.97)**
PSI scores	1.43 (1.22)	1.20 (0.68)	1.90 (1.80)**

[a] EDI = Eating Disorder Inventory. PSI = Physical Size Index. Values for "Photograph Ratings" denote the mean of the photograph ratings raw scores (range = 1, least physically-similar-5, most physically-similar) for each group.
* $p < .05$. The MZ twins are significantly more similar than the DZ twins.
** $p < .001$. The MZ twins are significantly more similar than the DZ twins.

Table II. Mean and (Standard Deviations) of Log Transformed EDI Absolute Difference Scores Stratified by Photograph Ratings[a]

	Rating 5	Rating 4	Rating 3	Rating 2	Rating 1
MZ twin pairs	*M* (*SD*) (n = 79)	*M* (*SD*) (n = 92)	*M* (*SD*) (n = 15)	—	—
EDI total score	3.49 (2.64)	4.04 (3.16)	4.73 (3.63)	—	—
Weight preoccupation	1.58 (1.50)	1.70 (1.66)	2.13 (1.30)	—	—
Body dissatisfaction	1.03 (1.23)	1.57 (1.42)	1.40 (1.35)	—	—
Compensatory behavior	0.42 (0.81)	0.49 (0.83)	0.53 (0.74)	—	—
Binge eating	1.10 (1.30)[b]	1.18 (1.22)	2.00 (1.20)	—	—
DZ twin pairs	*M* (*SD*) (n = 4)	*M* (*SD*) (n= 16)	*M* (*SD*) (n = 39)	*M* (*SD*) (n = 32)	*M* (*SD*) (n = 8)
EDI total score	5.25 (4.99)	4.31 (4.05)	6.32 (4.90)	5.45 (3.37)[c]	7.62 (6.35)
Weight. preoccupation	2.50 (1.91)	2.19 (1.42)	2.72 (2.05)	2.19 (1.71)	4.50 (2.62)
Body dissatisfaction	1.50 (1.29)	2.06 (1.69)	1.87 (1.79)	2.59 (1.79)	2.75 (2.05)
Compensatory behavior	0.50 (1.00)	0.38 (0.62)	0.60 (0.90)	0.52 (0.83)	1.25 (1.98)
Binge eating	1.25 (1.26)	1.48 (1.26)	1.85 (1.31)	1.56 (1.29)	1.00 (1.69)

[a] ED1 = Bating Disorder Inventory. Physical similarity increases with larger values of photograph ratings, from rating 1 (least similar, no physical resemblance) to rating S (most similar, twins are as similar as "two peas in a pod").
[b] n = 78.
[c] n = 33.

DZ ($F(4,99) = .22$, p = .93) twins for ED1 Total Score indicated minimal differences in general eating attitudes and behaviors among physically-less similar relative to more-similar twins. The nonsignificant MANOVAs for all of the ED1 subscales in both MZ ($\lambda = .92, F(8,358) = 1.87$, $p = .07$) and DZ ($\lambda = .82, F(16,270) = 1.16$, $p = .30$) twins paralleled these findings.

Photograph ratings were combined and independent sample t-tests, ANOVAs, and MANOVAs were performed to determine if observed nonsignificant results were due to insufficient samples sizes in groups when viewed independently. T-tests between rating 5 versus ratings 4 and 3 combined in MZ twins indicated no significant differences on any of the ED1 scales (all ps > .10), except Body Dissatisfaction; on this subscale, more physically-similar MZ twins had a lower mean within-pair difference score than less physically-similar MZ twins ($t(185) = -2.06$, $p = .04$).

Table III. Pearson Correlations Between PSI Absolute Difference Scores and Log Transformed ED1 Absolute Difference Scores'

	Correlations	
EDI scale	MZ twin pairs (n = 184486)	DZ twin pairs (n = 98400)
EDI total score	−.04	.06
Body dissatisfaction	−.11	.08
Compensatory behavior	−.15*	.08
Binge eating	.06	−.10
Weight preoccupation	−.07	.01

[a] EDI = Eating Disorder Inventory.
PSI= Physical Size Index.
* $p < 0.5$.

Combined ratings in DZ twins again yielded few significant differences. A one-way ANOVA for ED1 Total Score ($F(2,99) = .42$, $p = .66$) and a one-factor MANOVA ($\lambda = .92$, $F(8,180) = .94$, $p = .49$) for the ED1 subscales using combined ratings (Rating 5 and 4 combined, versus Rating 3, versus Rating 2 and 1 combined) showed no significant differences between physically-less and physically-more similar twins for eating attitudes and behaviors.

Finally, Pearson correlations between PSI and EDI within-twin pair absolute difference scores were relatively small and mainly nonsignificant (see Table III). In addition, regression analyses (data not shown) calculated for photograph ratings and PSI scores separately indicated that these measures accounted for little of the variance (range = 0.1%–4.0%) in ED1 within-pair absolute difference scores in both MZ and DZ twins.

DISCUSSION

The bulk of the evidence presented does not support the assumption that physical similarity accounts for increased MZ relative to DZ twin resemblance for eating attitudes and behaviors. Neither general physical similarity nor body size similarity showed significant associations with ED1 within-pair absolute difference scores. While some analyses displayed an overall pattern consistent with such an association, regression analyses suggested that physical similarity only accounted for nominal to minimal proportions of variance in eating attitudes and behaviors; Post-hoc power analyses indicated that the current study had adequate power (Average power: MZ twins = .89; DZ twins = .53) to detect mean differences of moderate magnitude (effect size ≥ SO) (Cohen, 1977), especially in the MZ twin sample. Assuming that a clinically-meaningful effect of physical similarity on twin resemblance for ED1 scores would be of moderate-to-high magnitude, it is unlikely that consequential effects went undetected in the current study. Nonetheless, results should be interpreted with caution, as the sample size in some of the most critical groups (i.e., physically-dissimilar MZ twins and physically-similar DZ twins) were relatively small. Overall, however, findings provide support for the EEA in twin studies of eating attitudes and behaviors.

The current investigation improved upon several methodological limitations of past research. Administering continuous measures of eating attitudes and behaviors increased statistical power. The use of a large, population-based twin sample decreased chances for confounding independent and dependent variables. Multiple measures of physical similarity were used, each with moderate-to-high reliability. Additionally, these measures were relatively independent of zygosity classification methods, as only one of five zygosity indices (the ponderal index, which uses height and weight measurements) overlapped with the physical similarity assessments. This relative independence of zygosity and similarity lead to an increased ability to separate the effects of these two variables.

Differences between current findings and those of Hettema *et al.* (1995), where significant associations were found, may speak to methodological differences in traits studied. Hettema *et al.* (1995) examined relationships between physical similarity and BN whereas the current study focused on eating attitudes and behaviors in a general twin population. Therefore, it may be that more severe forms of eating pathology do show such relationships. However, it could also be that physical similarity is confounded with clinical state in bulimic twins. Longitudinal designs assessing physical similarity prior to the onset of an eating disorder are needed to clarify the relationship between physical resemblance and the development of eating disorders in twins.

The current findings parallel those in other areas that have consistently shown a lack of relationship between physical similarity and twin similarity for personality (Matheny *et al.*, 1976; Plomin, 1976), cognition (Matheny *et* al., 1976), and other psychiatric disorders (Hettema *et* al., 1995). Similar to these studies, we did find greater physical resemblance among MZ relative to DZ pairs, which may create unequal environments for the two types of twins. However, findings thus far suggest that the unequal environments that may result

are not etiologically-related to the development of eating attitudes and behaviors as well as other behavioral characteristics in twins.

ACKNOWLEDGMENTS

Supported in part by NIH grants #DA05147 and AA09367, and NIMH Research Training Grant #MH17069.

REFERENCES

Bulik, C. M., Sullivan, P. F., and Kendler, K. S. (1998). Heritability of binge-eating and broadly defined bulimia nervosa. *Biol. Psychiat.* **44:1210–1218.**

Cohen, D. J., Dibble, E., Grawe, J. M., and Pollin, W. (1975). Reliably separating identical from fraternal twins. *Arch. Gen. Psychiat.* **32:1371–1375.**

Cohen, J. (1977). ***Statistical Power Analysis for the Behavioral Sciences - Revised,*** Academic Press, New York.

Fabsitz, R. R., Garrison, R. J., Feinleib, M., and Hjortland, M. (1978). A twin analysis of dietary intake: Evidence for a need to control for possible environmental differences in MZ and DZ twins. *Behav. Genet.* **8:15–25.**

Garner, D. M., Olmstead, M. P., and Polivy, J. (1983). Development and validation of a multidimensional eating disorder inventory for anorexia nervosa and bulimia. *Int. J. Eating Disord* 2: 15-34.

Hettema, J. M., Neale, M. C., and Kendler, K. S. (1995). Physical similarity and the equal-environments assumption in twin studies of psychiatric disorders. *Behav. Genet.* **25:327–335.**

Holland, A. J., Sicotte, N., and Treasure, J. (1988). Anorexia nervosa: Evidence for a genetic basis. *J. Psychosomat. Res. 32: 561-571.*

Iacono, W. G., Carlson, S. R., Taylor, J., Elkins, I. J., and McGue, M. (1999). Behavioral disinhibition and the development of substance use disorders: Findings from the Minnesota Twin Family Study. *Dev. Psychopathol.* **11:869–900.**

Iacono, W. G., Lykken, D. T., and McGue, M. (1996). Psychophysiological prediction of substance abuse. In Gordon, H. W., and Glantz, M. D. (eds.), ***Individual Differences in the Biobehavioral Etiology of Drug Abuse,*** NIDA Research Monograph 159, National Institute on Drug Abuse, Rockville, MD, pp. 129-149.

Kendler, K. S. (1983). Overview: Current perspective on twin studies of schizophrenia. *Am. J. Psychiat.* **140:1413–1425.**

Kendler, K. S., Heath, A. C., Neale, M., Kessler, R. C., and Eaves, L. J. (1993). A test of the equal-environment assumption in twin studies of psychiatric illness. *Behav. Genet.* **23:21–27.**

Kendler, K. S., MacLean, C., Neale, M., Kessler, R., Heath, A., and Eaves, L. (199 1). The genetic epidemiology of bulimia nervosa. *Am. J. Psychiat. 148:* **1627–1635.**

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., and Eaves, L. J. (1992a). A population based twin study of major depression in women: The impact of varying definitions of illness. *Arch. Gen. Psychiat.* **49:257–266.**

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., and Eaves, L. J. (1992b). The genetic epidemiology of phobias in women: The inter-relationship of agoraphobia, social phobia and simple phobia. *Arch. Gen. Psychiat.* **49:267–272.**

Kendler, K. S., Neale, M. C.. Kessler, R. C., Heath, A. C., and Eaves, L. J. (1992c). Generalized anxiety disorder in women: A population based twin study. *Arch. Gen. Psychiat.* **49:273–281.**

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., and Eaves, L. J. (19924). A population-based study of alcoholism in women. *J. Am. Med. Assoc. 268: 1877-1882.*

Klump, K. L., McGue, M., and Iacono, W. G. (in press). Age differences in genetic and environmental influences on eating attitudes and behaviors in female adolescent twins. *J. Abnorm. Psychol.*

Loehlin, J. C., and Nichols, R. C. (1976). ***Heredity, Environment and Personality,*** University of Texas Press, Austin, TX.

Lykken, D. T., Bouchard, T. J., McGue, M., and Tellegen, A. (1990). The Minnesota Twin Family Registry: Some initial findings. *Acta Gemellogicae Med.* **39:35–70.**

Matheny, A. P., Wilson, R. S., and Dolan, A. B. (1976). Relations between twins' similarity of appearance and behavioral similarity: Testing an assumption. *Behav. Genet.* **6:343–351.**

McGue, M., Lykken, D. T., and Iacono, W. G. (1996). Genotype-environment correlations and interactions in the etiology of substance abuse and related behaviors. In Gordon, H. W., and Glantz, M. D. (eds.), ***Individual Differences in the Biobehavioral Etiology of Drug Abuse,*** NIDA Research Monograph 159, National Institute on Drug Abuse, Rockville, MD, pp. 49-72.

Neale, M. C., and Cardon, L. R. (1992). ***Methodology for Genetic Studies of Twins and Families,*** Kluwer Academic Publisher: Boston, MA.

O'Connor, T. G., McGuire, S., Reiss, D., and Plomin, R. (1998). Co-occurrence of depressive symptoms and antisocial behavior in adolescence: A common genetic liability. ***J. Abnorm. Psychol.*** **107:27–37.**

Plomin, R. (1990). ***Nature and Nurture: An Introduction to Human Behavioral Genetics,*** Brooks/Cole Publishing Company, Pacific Grove, CA.

Plomin, R., DeFries, J. C., and McClearn, G. E. (1990). ***Behavioral Genetics: A Primer,*** 2nd Ed., W.H. Freeman Company, New York.

Plomin, R., Willerman, L., and Loehlin, J. C. (1976). Resemblance in appearance and the equal environments assumption in twin studies of personality traits. ***Behav. Genet.*** **6:43–52.**

Rutherford, J., McGuffin, P., Katz, R. J., and Murray, R. M. (19933 Genetic influences on eating attitudes in a normal female twin pair population. *Psychol. Med.* **23:425–436.**

Scarr, S., and Carter-Saltzman, L. (1979). Twin method: Defense of a critical assumption. ***Behav. Genet.*** **9:527–542.**

Sherman, D. K., Iacono, W. G., and Donnelly, J. M. (1995). Development and validation of body rating scales for adolescent females. ***Int. J. Eating Disord. 18: 1-7.***

Sullivan, P. F., Bulik, C. R., and Kendler, K. (1998). The genetic epidemiology of bingeing and vomiting. ***Brit. J. Psychiat*** **173:75–79.**

Wade, T., Martin, N. G., and Tiggeman, M. (1998). Genetic and environmental risk factors for the weight and shape concerns characteristic of bulimia nervosa. ***Psychol. Med.*** **28:76** ***I-771.***