Estrogen Moderates Genetic Influences on Binge Eating During Puberty:
Disruption of Normative Processes?

Kelly L. Klump, Natasha Fowler, Laura Mayhall, and Cheryl L. Sisk
Michigan State University

K. M. Culbert
University of Nevada, Las Vegas

S. Alexandra Burt
Michigan State University

Puberty is a critical period for changes in genetic effects for binge eating in girls. Previous twin studies show increases in genetic influences on binge eating from prepuberty (~0%) to midpuberty and beyond (~50%). However, little is known about the factors that drive these shifts in genetic effects. A small pilot study showed that pubertal activation of estrogen may contribute to increases in genetic influences, possibly via hormonally induced changes in gene expression. However, large-scale studies investigating hormone effects on genetic risk are lacking. Thus, the purpose of the present study was to examine the effects of estrogen on genetic influences for binge eating in 964 female twins (ages 8–16 years) from the Michigan State University Twin Registry. Binge eating was assessed with the Minnesota Eating Behaviors Survey, whereas afternoon saliva samples were assayed for estradiol levels using standard enzyme immunoassays. Twin moderation models showed substantial differences in genetic influences on binge eating across estradiol levels. Stronger genetic effects were observed at lower (rather than higher) estradiol levels, even when controlling for the effects of age, body mass index, the physical changes of puberty, and the onset of menses. Overall, findings suggest that comparatively lower levels of estradiol during this critical period may disrupt normative developmental processes and enhance genetic influences on binge eating.

General Scientific Summary
This study suggests that the activation of estrogen during pubertal development in girls may impact genetic risk for binge eating. Specifically, genetic influences on binge eating become more prominent in girls who have comparatively lower levels of estradiol during this critical period. These lower levels of estradiol may disrupt normative development and lead to increased risk for binge eating in later adolescence and adulthood.

Keywords: binge eating, estrogen, progesterone, genetic effects, twins

Supplemental materials: http://dx.doi.org/10.1037/abn0000352.supp

Binge eating is defined as the consumption of a large amount of food, in a short period of time, with a loss of control over the eating episode (American Psychiatric Association [APA], 2013). It is a core symptom of bulimia nervosa (BN) and binge eating disorder (BED) and is a common feature of many eating disorders (e.g., anorexia nervosa (AN) binge/purge type, other specified eating and feeding disorders (APA, 2013). Given the transdiagnostic nature of binge eating, understanding its etiology is likely to contribute to knowledge of risk factors for a range of eating pathology.

Puberty is a key phenotypic and genetic risk period for the development of binge eating (Hildebrandt, Alfano, Tricamo, & Pfaff, 2010; Klump, Culbert, O’Connor, Fowler, & Burt, 2017) and eating disorders (Harden, Kretsch, Moore, & Mendle, 2014; Young, 2010) in girls. Rates of binge eating are low in prepuberty but increase significantly in late-to-post puberty and into young adulthood in girls (Klump, 2013, for a review) and female animals (Klump, Culbert, & Sisk, 2017; Klump, Suisman, Culbert, Kashy, & Sisk, 2011). This pattern of phenotypic effects mimics changes in genetic influences on binge eating across puberty as well. The...
heritability of binge eating is essentially 0% in pre-early puberty but increases linearly across puberty to ~50% in late-post puberty and adulthood (Klump, Culbert, O’Connor, et al., 2017). An interesting aspect of these pubertal shifts is that they appear to be specific to girls. Rates of binge eating do not increase during puberty in boys (Klump, 2013), and the heritability of more general disordered eating symptoms remains constant across puberty in boys as well (Klump et al., 2012). Although recent data suggests that shifts in genetic risk for boys may occur during earlier developmental stages (e.g., during adrenarche; Culbert, Burt, & Klump, 2017), there appear to be sex-specific factors that substantially increase the risk for binge eating in girls during puberty.

One sex-specific factor that may contribute is estrogen. Increases in estrogen drive pubertal development in girls but not boys (Wilson, Foster, Kronenberg, & Larsen, 1998) and may contribute to the phenotypic and genetic differences described above. Estrogen increases linearly throughout puberty and maps on to the linear increases in genetic effects on binge eating observed previously. Although progesterone is also involved in pubertal development in girls, progesterone becomes activated much later during the pubertal period (i.e., after first ovulation; Wilson et al., 1998); an activation that is likely too late to be able to account for increases in heritability that are commonly observed around mid-puberty in girls (Culbert, Burt, McGue, Iacono, & Klump, 2009; Klump, Culbert, O’Connor, et al., 2017; Klump, McGue, & Iacono, 2003; Klump, Perkins, Alexandra Burt, McGue, & Iacono, 2007). Thus, estrogen has been the primary focus of pubertal theories of genetic risk (Hildebrandt et al., 2010; Klump, 2013; Klump, Culbert, O’Connor, et al., 2017; Klump, Keel, Sisk, & Burt, 2010; Klump et al., 2003; Young, 2010).

An additional reason for the focus on estrogen is its known genomic effects within the central nervous system (CNS). A critical function of estradiol during development is the regulation of gene transcription (Östlund, Keller, & Hurd, 2003). When activated by hormones, intracellular receptors act as transcription factors to regulate mRNA expression in neurobiological systems via activating or inhibiting target genes. Increased genetic effects on binge eating during puberty may therefore result from the activation of estradiol and the genomic effects of the hormone on the structure and/or function of the CNS (e.g., the production of neurotransmitters, their receptors, or their signal transduction mechanisms) or other biological systems (e.g., appetitive signals from the periphery).

Extant data indicate that estrogen regulates gene transcription within several neural systems (e.g., serotonergic, dopaminergic, opioidergic; Becker, 2009; Craft, 2008; Östlund et al., 2003) that are disrupted in binge eating and that could contribute to increases in genetic effects during puberty. Given the timing of pubertal increases in genetic effects (from early to midpuberty) versus increases in phenotypic rates of binge eating (late-to-post puberty), theories (Klump, 2013; Klump et al., 2017) have proposed that estrogen acts organizationally to impact binge eating, with estrogen-mediated genomic changes leading to (permanent) neuronal changes in brain structure/function during early midpuberty that later lead to increased phenotypic risk for binge eating in late adolescence and adulthood.

Unfortunately, it is difficult to test organizational hormone effects in humans, and the specific genes involved in binge eating remain unknown. Thus, as a first step, twin studies should examine differences/changes in heritability across puberty and hormonal milieus. Differences in heritability would reflect changes in the influence of genetic factors/genes. This approach would be a straightforward (and cost-effective) way to investigate whether estrogen may contribute to differences in genetic effects across puberty in girls.

Only one small pilot twin study has examined estrogen effects during puberty (Klump, Keel, et al., 2010). In this preliminary study, the degree of genetic influence on binge eating was significantly greater in twins with higher versus lower estradiol levels during puberty (Klump, Keel, et al., 2010). Although sample sizes were small (n = 98 pairs), the effects of estrogen were found to be independent of age, body mass index (BMI), and the physical changes of puberty, suggesting specific effects of estrogen on genetic influences for binge eating.

To date, there have been no large-scale studies of estrogen effects during puberty. Thus, the aim of the current study was to examine the effects of estrogen on genetic influences on binge eating during puberty in a sample of 964 preadolescent and adolescent female twins (ages 8–16). We took advantage of our larger sample size and conducted contemporary twin moderation models that allowed us to examine linear and nonlinear effects of estrogen on genetic influences while controlling for many confounding variables (i.e., age, BMI, physical changes of puberty, onset of menses) that could account for any estrogen moderation effects. We used a continuous measure of binge eating that has shown increases in genetic effects across pubertal development in this sample (Klump, Culbert, O’Connor, et al., 2017) and that is appropriate for use in girls across our broad age range. Finally, we directly ruled out the effects of progesterone (and Estrogen × Progesterone interactions) on genetic influences by exploring progesterone moderation in preliminary analyses.

**Method**

**Participants**

The sample consisted of 964 (MZ = 464 [48%]; DZ = 500 [52%]) female twins ages 8 to 16 (M = 11.75, SD = 2.03) who participated in the Twin Study of Mood, Behavior, and Hormones during Puberty (TSMBH; O’Connor, Burt, VanHuyse, & Klump, 2016) from within the population-based, Michigan State University Twin Registry (MSUTR; for MSUTR and recruitment details, Burt and Klump (2013a) and Klump and Burt (2006)). The MSUTR recruits twins using birth records in collaboration with the Michigan Department of Health and Human Services (see Klump & Burt, 2006 and Burt & Klump, 2013b for recruitment details), and then the TSMBH (and other follow-up studies—Burt, Klump, Gorman-Smith, & Neiderhiser, 2016; Klump et al., 2013) recruits twins directly from the MSUTR pool. Response rates for the MSUTR (Burt & Klump, 2013a) and the TSMBH (~65%) are quite good and are on par or better than response rates for other large-scale twin registries (Baker, Barton, & Raine, 2002; Hay, McStephen, Levy, & Pearssall-Jones, 2002). MSUTR twins are representative of the recruitment region in terms of race, ethnicity, and socioeconomic status (Burt & Klump, 2013a).
Because the TSMBH was focused on examining hormone regulation during puberty, twins were required to meet the following inclusion/exclusion criteria to ensure the validity/reliability of the hormone samples: (a) no hormonal contraceptive use within the past 3 months; (2) no psychotropic or steroid medications within the past 4 weeks; (3) no pregnancy or lactation within the past 6 months; and (4) no history of genetic or medical conditions known to influence hormone functioning or appetite/weight. Despite these criteria, the TSMBH twins are representative of the recruitment region in terms of race/ethnicity (81.4% White, 7.9% Black, 0.6% Asian/Pacific Islander, 0.2% Native American, 7.7% multiracial, 2.1% missing data; 94.4% Non-Hispanic, 4.0% Hispanic, 1.7% missing data), and they do not differ from other MSUTR twins on measures of binge eating (Cohen’s $d = .12$). This study was approved by the Michigan State University Institutional Review Board (protocol #01-052M).

Measures

**Zygosity determination.** Zygosity was determined using a well-validated physical similarity questionnaire that has been shown to be over 95% accurate when compared to genotyping (Lykken, Bouchard, McGue, & Tellegen, 1990; Peeters, Van Geertel, Vlietinck, Derom, & Derom, 1998). The twins’ mother and trained research assistants independently completed the questionnaire for all twins. Reports were compared among raters, and discrepancies were resolved using questionnaire responses, review of photographs of the twins by the principal investigator (Kelly L. Klump), and twin concordance across several nucleotide polymorphisms.

**Binge eating.** Binge eating was assessed using the seven-item binge eating scale (assessing the tendency to engage in binge eating and to think about binge eating) from the Minnesota Eating Behavior Survey$^1$ (von Ranson, Klump, Iacono, & McGue, 2005). The MEBS binge eating scale is appropriate for use in prepubertal twins who self-reported one or more episodes of objective binge eating (on the Eating Disorders Examination Questionnaire; Fairburn & Beglin, 1994) scored significantly higher on the MEBS binge eating scale than controls, $t(992) = 15.67, p < .0001; d = 1.02$. Pearson correlations with related phenotypes (e.g., BMI, levels of body dissatisfaction/weight preoccupation, etc.) were significant and nearly identical across hormone levels as well (see Supplemental Table S1 in the online supplemental material).

As noted above, the MEBS has been employed in numerous developmental twin studies of eating disorders (Culbert et al., 2009; Klump, Burt, et al., 2007; Klump, Culbert, O’Connor, et al., 2017; Klump, Keel, et al., 2010; Klump et al., 2003; Klump, McGue, & Iacono, 2000; Klump, Perkins, et al., 2007), including a recent study of pubertal differences in genetic influences in the current TSMBH sample (Klump, Culbert, O’Connor, et al., 2017). Our use of the same, developmentally appropriate measure allowed us to explore whether estrogen accounts for previously observed pubertal differences in genetic effects.

**Hormone levels.** Salivary samples were used to obtain estradiol and progesterone levels. Saliva samples are preferred over other methods (e.g., blood spots) because they represent a less invasive collection method, particularly for preadolescent and adolescent participants who, in our lab studies, are less likely to agree to blood draws. Moreover, previous research has shown that saliva samples are associated with higher compliance and more robust hormone-behavior associations than blood-spot sampling (Edler, Lipson, & Keel, 2007). Finally, estradiol values from matched serum and saliva samples show a strong linear relationship for females ($r = .80$; Shirtcliff et al., 2000).

Following previously reported procedures (Klump, Keel, et al., 2010), twins were asked to refrain from eating or drinking for four hours prior to providing a saliva sample and to refrain from smoking, brushing their teeth, or chewing gum for 30 min prior to the saliva collection. Twins were then instructed to passively drool down a straw into a cryovial until ≥4 ml was produced. All saliva collections occurred between 2:00 p.m. and 5:00 p.m. We chose to collect during these times because afternoon–early evening diurnal variations in ovarian hormones during puberty tend to be minimal (Angold, Costello, Erkanli, & Worthman, 1999; Grumbach & Styne, 1998), and most families were available for assessments during this time. Saliva samples were frozen immediately and stored until transport.

We collected saliva samples on three consecutive days: Two samples were collected by the twins in their home, while the third was collected in our MSU laboratory during the primary, in-person lab assessment. Unfortunately, hormone values were less correlated ($r’s = .53–.58$) across the assessment days than would be expected. Reasons for this are unclear, although it seems likely that twin difficulties adhering to our saliva collection procedures (e.g., collection between 2 and 5 p.m., fasting for 4 hr, immediately freezing samples and keeping them frozen during transport) likely contributed. Thus, in the current report, we focus on the in-person assessment hormone values only as they were collected on the same day as the binge eating scores, and our in-lab collection procedures were carefully standardized and minimized error. However, our results were similar when using the average of the three assessment days (see Supplemental Figure S1 in the online sup-

---

$^1$The Minnesota Eating Behavior Survey (MEBS; previously known as the Minnesota Eating Disorder Inventory) was adapted and reproduced by special permission of Psychological Assessment Resources, Inc., 16204 North Florida Avenue, Lutz, FL 33549, from the Eating Disorder Inventory (collectively, EDI and EDI-2) by Garner, Olmstead, and Polivy (1983) by the Psychological Assessment Resources, Inc. Further reproduction of the MEBS is prohibited without prior permission from Psychological Assessment Resources, Inc.
plemental material), suggesting that our findings are robust to saliva collection procedures and sampling frequency.

All saliva samples were processed in duplicate by Salimetrics, LLC (State College, PA) using enzymatic immunoenzyme kits designed specifically for analyzing saliva. These assays show excellent intra- and inter-assay coefficients of variation (estradiol = 7.1% and 7.5%; progesterone = 6.2% and 7.6%), as well as assay sensitivity (measured by interpolating the mean optical density minus 2 SDs of 10–20 replicates at the 0 pg/ml level; estradiol = 0.10 pg/ml; progesterone = 5 pg/ml) and method accuracy (determined by spike recovery and linearity; estradiol = 104.2% and 99.4%; progesterone = 99.6% and 91.8%). Notably, all of the twins in the current study had estradiol and progesterone levels that fell above the minimum detection limits for the assays.

**Physical changes of puberty.** The physical changes of puberty were measured with twin self-report on the Pubertal Development Scale (PDS; Petersen, Crockett, Richards, & Boxer, 1988). The PDS assesses growth spurts, skin changes, body hair growth, and breast development on a 4-point scale, ranging from 1 (development has not yet begun) to 4 (development seems complete). Menses was rated as either present (coded 1) or absent (coded 4). Mernal reports on the PDS were used for a subset of twins (n = 16; 1% of sample) who were missing self-reported PDS scores. Consistent with past studies (Culbert et al., 2009; Klump et al., 2003), PDS items were summed and averaged for all analyses. Previous studies of the PDS psychometric properties have supported its reliability (median α = .77; Petersen et al., 1988) and validity, including showing high correlations with clinician ratings of these same physical characteristics (r’s = .61–.67; Petersen et al., 1988).

**BMI.** BMI was calculated (weight in kg/height in m²) using height and weight measured with a wall-mounted ruler and digital scale, respectively. We used raw (instead of age- and gender-adjusted) BMIs in analyses to be consistent with our recent study examining differences in genetic effects across pubertal development in binge eating (Klump, Culbert, O’Connor, et al., 2017) and prior developmental twin studies (Culbert et al., 2009; Culbert, Burt, Sisk, Nigg, & Klump, 2014; Klump et al., 2003, 2012; Klump, Burt, et al., 2007; Klump et al., 2010; Klump, Holly, Iacono, McGuie, & Willson, 2000; Klump, Keel, et al., 2010). Raw and age-adjusted BMI scores were nearly identical (r > .99; M difference = 0.04, SD = 0.03).

**Statistical Analyses.**

**Data preparation.** Missing scores on the binge eating scale were prorated when ≤10% of the items were missing, and coded missing if >10% of items were missing. These scores were then log transformed prior to analyses to account for positive skew. We conducted all analyses with the covariates age, BMI, and PDS scores regressed out of each twin’s binge eating score to ensure that differences in etiologic effects were specific to hormone levels rather than these potentially confounding variables. Notably, however, results were similar when using raw binge eating scores that did not partial out these variables (see Supplemental Figure S2 in the supplemental online material), suggesting that they had minimal impact on differences in genetic effects across hormone levels during puberty.

**Twin correlations.** Twin intraclass correlations were used as initial indicators of differences in genetic and environmental influences on binge eating across estradiol levels. Because MZ twins share approximately 100% of their segregating genes while DZ twins share, on average, 50%, significantly greater MZ relative to DZ twin correlations indicate the presence of additive genetic effects (i.e., effects that add or sum across genes). By contrast, MZ and DZ twin correlations that are similar in magnitude and are significantly greater than 0 signify a lack of genetic effects, but significant shared environmental influences (i.e., environmental factors that are common to siblings growing up in the same family and contribute to their behavioral similarity). Finally, MZ twin correlations that are less than 1.00 indicate the presence of non-shared environmental factors (i.e., factors that are unique to siblings growing up in the same family and contribute to behavioral differences) and measurement error.

Although continuous estradiol levels were used in model-fitting (see below), a dichotomous estrogen group (i.e., low/high) was developed for the twin correlations based on median splits for estradiol levels (cut-off = 1.26 pg/ml). Only pairs concordant on estrogen group (e.g., both cotwins had low estradiol levels) could be included in correlations (N = 534 twins from 267 pairs; 55% of the total sample), although concordant and discordant cotwins were included in model-fitting analyses (see below).

**Twin moderation models.** We used extended, univariate twin moderation models (see Figure 1; van der Sluis, Posthuma, & Dolan, 2012) to examine differences in additive genetic (A), shared environmental (C), and nonshared environmental (E) influences on binge eating across estradiol levels. These models estimate (a) path coefficients (i.e., a, c, e) assessing genetic/environmental influences at the lowest level of estradiol; (b) linear moderators (βXM, βYM, βEM) assessing linear increases/decreases in etiologic influences across estradiol levels; and (c) quadratic moderators (βXM², βYM², βEM²) assessing nonlinear increases/decreases in genetic/environmental influences across estradiol levels.

We first fit the “full” model that included all parameters. We then compared the fit of this model to nested submodels that differentially constrained the linear and quadratic genetic, shared environmental, and nonshared environmental moderaters to 0. We initially focused on submodels that are commonly tested in twin moderation analyses (i.e., models that constrained all moderators to 0, models that constrained the quadratic moderators only to 0, etc.) as well as those that directly tested our theory of genetic moderation (i.e., models that constrained the genetic moderators to 0). Because of the large number of submodels that could be fit, however, we used parameter estimates from the full model to identify additional submodels to fit. This approach allowed us to test relevant submodels without unduly increasing the number of tests.

We used the raw estradiol values in all analyses, although results were nearly identical when “binning” the estradiol values into two groups (see Supplemental Figure S3 in supplemental material). Prior to model-fitting, binge eating scores were standardized and minimum estradiol levels were “floored” to 0. Models were then fit to the raw data using the maximum likelihood option in Mx (Neale, Boker, Xie, & Maes, 2003). This option treats missing data as missing-at-random and is expected to produce less biased and more consistent estimates than other techniques (e.g., listwise
Model fit comparisons were made by taking the difference in minus twice the log-likelihood (−2lnL) between the full and nested models, which is chi-squared distributed under the null hypothesis implied by the reduced model. Large (statistically significant) differences led to a rejection of the nested model. Akaike’s information criterion (AIC), Bayesian information criterion (BIC), sample-size adjusted BIC, and deviance information criterion (DIC) were also used to select the best-fitting model, where models that minimized the largest number of these scores were preferred. Unstandardized parameter estimates from the full and best-fitting models are reported in figures, as they more accurately depict absolute differences in etiologic effects than standardized estimates that represent differences as proportions of the total variance. However, we report standardized estimates in the text, as these estimates allow for a direct comparison of our findings to previous studies.

Notably, the twin moderation models described above rest on one important assumption—that the moderator (i.e., estradiol levels) is genetically independent of the dependent variable (i.e., binge eating). If there is not independence, then genetic mediation (i.e., termed gene-environment correlations or rGE) may be present for estradiol levels and binge eating, such that the same genes that influence hormone levels also influence binge eating. These types of rGE are troublesome in the context of twin moderation models, as rGE could conceivably “masquerade” as hormone moderation effects. For example, if binge eating stems, in part, from genes common to hormone levels, then increases in genetic influences at high levels of hormones could be a reflection of shared genes and/or environments, rather than an independent effect of the hormone levels on genetic influences on binge eating. To avoid this interpretive difficulty, twin researchers are advised to use the “gene × environment (GxE) in the presence of rGE” model (Purcell, 2002) that tests the significance of genetic and/or environmental mediation by comparing the fit of models that allow for genetic/environmental covariance between the moderator and the dependent variable to models that constrain the covariance to be zero. Consequently, before fitting the models described above, we fit “GxE in the presence of rGE” models and found that the moderation of the covariance between estrogen and binge eating was nonsignificant (i.e., the parameter could be constrained to 0 without a worsening of model fit—comparison with full model: \( \chi^2(3) = 18.45, p = .0004 \)), outside of any rGE processes. These findings provided the first indication that our twin models would show significant effects of estrogen on genetic and/or environmental risk for binge eating that reflect independent hormone effects rather than rGE.

Finally, as noted in the introduction, we also conducted preliminary model-fitting analyses to confirm that progesterone did not significantly moderate etiologic effects on binge eating during puberty. We examined these effects using two-moderator models that allowed us to examine whether estrogen alone, progesterone alone, or Estrogen × Progesterone interactions moderated etiologic effects on binge eating. Importantly, the best-fitting model constrained all of the progesterone moderators (including the Estrogen × Progesterone interaction) to 0 (i.e., comparison with full model: \( \chi^2(6) = 5.79, p = .45 \)), but retained the estrogen moderators (see Supplemental Figure S3 in supplemental online material). These preliminary results supported our theories of estrogen (rather than progesterone) moderation (Klump, 2013; Klump et al., 2017) and our focus on estrogen in the primary analyses.

Results

Descriptive Statistics and Phenotypic Correlations

Descriptive statistics and intercorrelations among our study variables are presented in Table 1. A range of binge eating was present in our sample, as scores varied across the full spectrum of possible
scores (i.e., 0–7), and 3.2% of twins scored above the clinical cut-off for binge eating (i.e., clinical cut-off = the mean of individuals with bulimia nervosa score ≥4.36; von Ranson et al., 2005). Inter correlations among the study variables were as expected with binge eating showing significant (albeit modest) correlations with PDS scores and BMI, and estradiol levels showing significant and positive correlations with age, PDS scores, and BMI as well. Binge eating scores did not correlate significantly with estradiol levels, a finding that has been reported previously in preadolescent/adolescent girls (Klump, Keel, et al., 2010). Importantly, however, the lack of significant phenotypic associations does not negate a moderating effect of the hormone on genetic effects, as the presence of genetic moderation would be expected to attenuate phenotypic associations. This pattern has been observed previously in twin studies, where pubertal development was a significant moderator of genetic effects despite modest phenotypic associations with binge eating or other types of disordered eating (Culbert et al., 2009; Klump, Culbert, O’Connor, et al., 2017; Klump, Keel, et al., 2010; Klump et al., 2003; Klump, Perkins, et al., 2007).

**Twin Correlations**

Twin correlations suggested significant moderation of genetic effects by estradiol levels (see Table 2), but the pattern of effects was opposite to that found previously in pilot work (Klump, Keel, et al., 2010). Lower levels of estradiol were associated with stronger genetic influences, as the MZ twin correlation was significantly greater than the DZ twin correlation, and the MZ/DZ difference was of large effect size (see q estimates in Table 2). By contrast, no significant MZ/DZ twin differences were observed at higher estradiol levels (and effect sizes were small—see Table 2), suggesting minimal genetic influences but possible shared environmental effects. Nonshared environmental influences were present across all estradiol levels, as the MZ twin correlation was less than 1.00, but seemingly increased at higher estradiol levels where the MZ twin correlation was quite small. Notably, in the higher estradiol group, the DZ twin correlation was larger than the MZ twin correlation, but the MZ/DZ correlation difference was nonsignificant and of small effect size. Overall then, findings from the twin correlations seemed to show stronger genetic effects on binge eating at lower versus higher estradiol levels, and stronger nonshared environmental influences as estradiol levels increased.

**Twin Moderation Models**

Twin model-fitting analyses confirmed these impressions. Examination of the parameter estimates from the full model (see Figure 2a) suggested significant (and linear) decreases in genetic influences across estradiol levels, quadratic/nonlinear changes in nonshared environmental effects, but minimal differences in shared environmental effects. Model fit comparisons (see Table 3) indicated that the model constraining all of the shared environmental moderators (linear and quadratic) to 0 provided a good fit to the data (i.e., a nonsignificant chi-square change and lower AIC, BIC, sample-adjusted BIC, and DIC values than most models—see Table 3), whereas models constraining all of the genetic, shared environmental, and nonshared environmental moderators to 0 provided a poorer fit (i.e., significant chi-square change tests and higher values for some fit indices—see Table 3). Importantly, models constraining the genetic moderators

<table>
<thead>
<tr>
<th>Variable</th>
<th>Range</th>
<th>M (SD)</th>
<th>Binge eating</th>
<th>Estradiol</th>
<th>Age</th>
<th>PDS</th>
<th>BMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binge eating</td>
<td>0–7</td>
<td>.90 (1.34)</td>
<td>1.00</td>
<td>–.01</td>
<td>.04</td>
<td>.11**</td>
<td>.13***</td>
</tr>
<tr>
<td>Estradiol (pg/mL)</td>
<td>.11–4.69</td>
<td>1.43 (1.82)</td>
<td>–</td>
<td>1.00</td>
<td>.24***</td>
<td>.23***</td>
<td>.22**</td>
</tr>
<tr>
<td>Age</td>
<td>8–16</td>
<td>11.75 (2.03)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>.80***</td>
<td>.41***</td>
</tr>
<tr>
<td>PDS score</td>
<td>1–4</td>
<td>2.25 (3.0)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1.00</td>
<td>.49**</td>
</tr>
<tr>
<td>BMI</td>
<td>10.70–46.55</td>
<td>19.53 (4.47)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* PDS = Pubertal Development Scale; BMI = body mass index. Binge eating scores were log transformed prior to analysis, but raw values are presented in the table for descriptive purposes. **p < .01. ***p < .001.

---

**Table 1**

**Descriptive Statistics and Intercorrelations for Study Variables (N = 964 Twins)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Range</th>
<th>M (SD)</th>
<th>Binge eating</th>
<th>Estradiol</th>
<th>Age</th>
<th>PDS</th>
<th>BMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binge eating</td>
<td>0–7</td>
<td>.90 (1.34)</td>
<td>1.00</td>
<td>–.01</td>
<td>.04</td>
<td>.11**</td>
<td>.13***</td>
</tr>
<tr>
<td>Estradiol (pg/mL)</td>
<td>.11–4.69</td>
<td>1.43 (1.82)</td>
<td>–</td>
<td>1.00</td>
<td>.24***</td>
<td>.23***</td>
<td>.22**</td>
</tr>
<tr>
<td>Age</td>
<td>8–16</td>
<td>11.75 (2.03)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>.80***</td>
<td>.41***</td>
</tr>
<tr>
<td>PDS score</td>
<td>1–4</td>
<td>2.25 (3.0)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1.00</td>
<td>.49**</td>
</tr>
<tr>
<td>BMI</td>
<td>10.70–46.55</td>
<td>19.53 (4.47)</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2**

**Twin Correlations for Binge Eating by Estradiol Levels**

<table>
<thead>
<tr>
<th>Estrogen groups</th>
<th>MZ pairs</th>
<th>DZ pairs</th>
<th>Z test of independence</th>
<th>p</th>
<th>q</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full sample (N = 267 pairs)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower estradiol (MZ n = 67 pairs; DZ n = 62 pairs)</td>
<td>.70*** (.55, .80)</td>
<td>.10 (.15, .34)</td>
<td>4.25</td>
<td>&lt;.001</td>
<td>.77</td>
</tr>
<tr>
<td>Higher estradiol (MZ n = 69 pairs; DZ n = 69 pairs)</td>
<td>.11 (.13, .34)</td>
<td>.27 (.04, .48)</td>
<td>.96</td>
<td>.34</td>
<td>.16</td>
</tr>
<tr>
<td>Premenarcheal twins only (n = 141 pairs, 282 twins)</td>
<td>.74*** (.55, .86)</td>
<td>.50** (.18, .73)</td>
<td>1.59</td>
<td>.11</td>
<td>.40</td>
</tr>
<tr>
<td>Lower estradiol (MZ n = 59 pairs; DZ n = 51 pairs)</td>
<td>.001 (.33, .33)</td>
<td>.22 (.12, .52)</td>
<td>.90</td>
<td>.37</td>
<td>.22</td>
</tr>
</tbody>
</table>

*Note.* MZ = monozygotic; DZ = dizygotic; Z test of independence = a Z test of significant differences between the MZ and DZ twin correlations; q = effect size for the difference between the two twin correlations (no effect q < .10, small effect q = .10–.30, medium effect q = .30–.50, large effect q > .50). The 95% confidence intervals are presented in parentheses. Binge eating scores were log transformed prior to analysis. **p < .01. ***p < .001. The twin correlation is significantly different from zero.
(linear and quadratic) only, and the nonshared environmental moderators (linear and quadratic only) only, also resulted in a significant worsening of model fit (significant chi-square change tests and higher values for fit indices—see Table 3). Overall, these initial model fit comparisons confirmed that there is significant moderation of genetic effects (and nonshared environmental influences) but no significant shared environmental moderation by estradiol levels during puberty.

In terms of additional submodels to fit, parameter estimates from the full model seemed to show linear decreases in genetic, but nonlinear differences in nonshared environmental effects, across estradiol levels. We therefore fit three additional submodels that constrained all of the shared environmental moderators to 0, but then differentially constrained the quadratic genetic moderator, quadratic nonshared environmental moderator, or both to 0 (see Table 3). Importantly, the benefit of these models is that they directly tested whether the significant differences in genetic and nonshared environmental influences across estradiol levels were linear (i.e., the quadratic moderator could be constrained to 0) or nonlinear in nature (i.e., the quadratic moderator could not be constrained without a worsening of model fit). Two of these models (i.e., the model that constrained the genetic quadratic moderator only and the model that constrained both the genetic and nonshared environmental quadratic moderators) showed essentially equivocal evidence for being the best-fitting model. Both models exhibited nonsignificant chi-square change values, but two out of the four fit indices favored one model over the other (i.e., the AIC and sample-adjusted BIC favored the model constraining the genetic quadratic moderator only, while the BIC and DIC favored the model constraining both the genetic and nonshared environmental quadratic moderators—see Table 3). We took the more conservative approach and chose the model that retained the nonshared environmental quadratic moderator (i.e., the model that constrained the genetic quadratic model only) as best fitting, as the nonshared environmental quadratic moderator was statistically significant (i.e., 95% confidence interval did not include 0—see the footnote in Table 3) in this model. Unstandardized (see Figure 2a) and standardized (see parentheses below) estimates from this final model confirmed the presence of

![Figure 2. Unstandardized Estimates of Additive Genetic (a), Shared Environmental (c), and Nonshared Environmental (e) Effects from the Full and Best-Fitting Models in a) All Twins (N = 964); and b) Premenarcheal Twins Only (N = 530). For the full sample, the best-fitting model was the model that constrained all of the shared environmental moderators to 0 as well as the genetic quadratic moderator. The best fitting model for premenarcheal twins was the model that constrained all of the shared environmental moderators as well as the genetic and nonshared environmental quadratic moderators to 0. It is important to note that the estradiol values in the figures were floored to 0 for the twin models (see Method), and that the raw estradiol values range from 0.11 to 4.69 in the full sample, and 0.11 to 4.56 in the sample of premenarcheal twins only.]
ESTROGEN EFFECTS

465

Table 3
Model Fit Comparisons for the Twin Moderation Models

<table>
<thead>
<tr>
<th>Model</th>
<th>$-2\ln L (df)$</th>
<th>$\chi^2 (df)$</th>
<th>$p$</th>
<th>AIC</th>
<th>BIC</th>
<th>Adj. BIC</th>
<th>DIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full sample ($n = 964$ twins)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full model</td>
<td>2227.64 (780)</td>
<td>—</td>
<td>—</td>
<td>667.64</td>
<td>—1226.73</td>
<td>10.78</td>
<td>—509.96</td>
</tr>
<tr>
<td>Nested Submodels:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constrain A, C, E, quad mods</td>
<td>2232.75 (783)</td>
<td>5.11 (3)</td>
<td>.16</td>
<td>666.75</td>
<td>—1233.18</td>
<td>9.09</td>
<td>—513.65</td>
</tr>
<tr>
<td>Constrain linear and quad A mods</td>
<td>2255.69 (782)</td>
<td>28.05 (2)</td>
<td>&lt;.001</td>
<td>691.69</td>
<td>—1218.71</td>
<td>21.98</td>
<td>—500.10</td>
</tr>
<tr>
<td>Constrain linear and quad C mods</td>
<td>2228.25 (782)</td>
<td>.61 (2)</td>
<td>.74</td>
<td>664.25</td>
<td>—1232.43</td>
<td>8.26</td>
<td>—513.82</td>
</tr>
<tr>
<td>Constrain all C mods and A quad only</td>
<td>2229.97 (783)</td>
<td>2.33 (3)</td>
<td>.51</td>
<td>663.97</td>
<td>—1234.57</td>
<td>7.71</td>
<td>—515.04</td>
</tr>
<tr>
<td>Constrain all C mods and E quad only</td>
<td>2232.68 (783)</td>
<td>5.04 (3)</td>
<td>.17</td>
<td>666.88</td>
<td>—1233.21</td>
<td>9.06</td>
<td>—513.68</td>
</tr>
<tr>
<td>Constrain all C mods, A, E quad</td>
<td>2272.82 (784)</td>
<td>5.18 (4)</td>
<td>.27</td>
<td>664.82</td>
<td>—1236.14</td>
<td>7.72</td>
<td>—515.70</td>
</tr>
<tr>
<td>Constrain linear and quad E mods</td>
<td>2241.62 (782)</td>
<td>13.98 (2)</td>
<td>&lt;.001</td>
<td>677.62</td>
<td>—1225.74</td>
<td>14.95</td>
<td>—507.13</td>
</tr>
<tr>
<td>Constrain all linear and quad moderators for A, C, and E</td>
<td>2242.05 (786)</td>
<td>14.41 (6)</td>
<td>.03</td>
<td>670.05</td>
<td>—1237.53</td>
<td>9.50</td>
<td>—515.25</td>
</tr>
<tr>
<td>Premenarcheal twins only ($n = 530$ twins)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full model</td>
<td>1420.77 (510)</td>
<td>—</td>
<td>—</td>
<td>400.77</td>
<td>—713.50</td>
<td>95.09</td>
<td>—244.75</td>
</tr>
<tr>
<td>Nested submodels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constrain A, C, E, quad mods</td>
<td>1421.80 (513)</td>
<td>1.03 (3)</td>
<td>.79</td>
<td>395.80</td>
<td>—721.27</td>
<td>91.98</td>
<td>—249.85</td>
</tr>
<tr>
<td>Constrain linear and quad A mods</td>
<td>1423.33 (512)</td>
<td>2.56 (2)</td>
<td>.28</td>
<td>390.33</td>
<td>—717.71</td>
<td>93.95</td>
<td>—247.22</td>
</tr>
<tr>
<td>Constrain linear and quad C mods</td>
<td>1420.97 (515)</td>
<td>.20 (2)</td>
<td>.90</td>
<td>396.97</td>
<td>—718.89</td>
<td>92.77</td>
<td>—248.40</td>
</tr>
<tr>
<td>Constrain all C mods, A quad only</td>
<td>1421.62 (513)</td>
<td>.85 (3)</td>
<td>.84</td>
<td>395.62</td>
<td>—721.36</td>
<td>91.89</td>
<td>—249.94</td>
</tr>
<tr>
<td>Constrain all C mods, E quad only</td>
<td>1421.63 (513)</td>
<td>.96 (3)</td>
<td>.81</td>
<td>395.73</td>
<td>—721.30</td>
<td>91.95</td>
<td>—249.89</td>
</tr>
<tr>
<td>Constrain all C mods, A, E quad</td>
<td>1422.05 (514)</td>
<td>1.28 (4)</td>
<td>.86</td>
<td>394.05</td>
<td>—723.93</td>
<td>90.90</td>
<td>—251.69</td>
</tr>
<tr>
<td>Constrain linear and quad E mods</td>
<td>1431.66 (512)</td>
<td>10.89 (2)</td>
<td>.004</td>
<td>407.66</td>
<td>—713.55</td>
<td>98.12</td>
<td>—243.05</td>
</tr>
<tr>
<td>Constrain A, C, Moderators</td>
<td>1433.49 (516)</td>
<td>12.72 (6)</td>
<td>.047</td>
<td>401.49</td>
<td>—728.80</td>
<td>94.21</td>
<td>—249.63</td>
</tr>
</tbody>
</table>

Note. $-2\ln L = \text{minus twice the log-likelihood}; \chi^2 = \chi^2$-square change; AIC = Akaike information criterion; BIC = Bayesian information criterion; Adj. BIC = sample size adjusted BIC; DIC = deviance information criterion; Full model = model with paths and all moderators; mods = moderators; quad = quadratic or non-linear moderators; A = additive genetic effects; C = shared environmental effects; E = nonshared environmental effects. Each nested model is compared to the full model when calculating the $\chi^2$ and degrees of freedom. The best-fitting models are noted by bolded text. Unstandardized estimates (with 95% confidence intervals in parentheses) from the best-fitting models were as follows: full sample: $a = -.81 (-1.04, -.49)$, $c = .36 (-.59, .59)$, $e = -.42 (-.64, -.26)$, $\beta_2M = .21 (.02, .43)$, $\beta_2M = .21 (-.58, .11)$, $\beta_3M^2 = 0$, $\beta_4M^2 = 0$, $\beta_5M^2 = .05 (-.01, .12)$; Premenarcheal twins only: $a = .78 (.34, 1.06)$, $c = -.41 (-.62, .62)$, $e = -.48 (-.66, -.33)$, $\beta_2M = -.26 (-.48, -.04)$, $\beta_3M = 0$, $\beta_4M = -.20 (-.31, -.09)$, $\beta_5M^2 = 0$, $\beta_6M^2 = 0$, $\beta_7M^2 = 0$.

substantial differences in genetic and nonshared environmental influences across estradiol levels, with significantly greater genetic effects at lower (69%) versus higher (2%) estradiol levels, and significantly greater nonshared environmental influences at higher (68%) versus lower (18%) estradiol levels. Shared environmental influences remained stable (13%) and nonsignificant (i.e., 95% confidence intervals included 0—see Table 3’s footnote) across all estradiol levels.

Because findings for genetic effects were opposite to those of our small pilot study (Klump, Keel, et al., 2010), we conducted post hoc analyses examining only those twins who were premenarcheal to determine if our inclusion of pre- and postmenarcheal twins may have influenced our results. Our pilot study also included pre/postmenarcheal twins, but the age range of the current sample was much broader and included a larger proportion of postmenarcheal twins (Klump, Keel, et al., 2010). Hormone values would be expected to be more variable in postmenarcheal twins and dependent upon menstrual cycle phase. In the current study, we attempted to standardize the day of the lab assessment for our menstruating twins (i.e., Day 15 of the menstrual cycle) to ensure more stable hormone levels, but this was not always possible given twin scheduling constraints and the error associated with relying on twin self-reports for estimating cycle days.

Twin correlations in the premenarcheal only group (see Table 2) showed a very similar pattern of results to the full sample (i.e., stronger genetic influences at lower as compared to higher estradiol levels). The only exception was that the MZ/DZ twin correlation differences for high estradiol levels was no longer significant (but still represented a medium effect size (q)—see Table 2). Model-fitting results in premenarcheal twins were also very similar to the full sample, with stronger genetic influences at lower estradiol levels, stronger nonshared environmental effects at higher estradiol levels (see Figure 2b), and minimal differences in shared environmental influences. In this smaller sample, the best-fitting model constrained all shared environmental moderators and the quadratic genetic moderator to 0 (like in the full sample) and also constrained the quadratic nonshared environmental moderator to 0. Unstandardized (see Figure 2) and standardized (see parentheses) estimates from the best-fitting model once again showed stronger genetic effects at lower (60%) versus higher (7%) estradiol levels, significantly greater nonshared environmental influences at higher (85%) versus lower (23%) estradiol levels, and minimal/nonsignificant shared environmental influences (18%).

Discussion

This was the first large-scale study of the effects of estrogen on genetic influences on binge eating during puberty in girls. Estradiol levels significantly moderated genetic and environmental in-

As an additional check, we also fit a series of two moderator models (i.e., Estrogen X Menses) examining whether etiologic effects differed in pre- versus postmenarcheal twins. The menses moderator could be constrained to 0 without a significant worsening of model fit $\chi^2(6) = 8.55$, $p = .20$, see Supplemental Figure S4 in the online supplemental material), suggesting that there were no statistically significant differences in etiologic effects between premenarcheal and postmenarcheal twins.
fluences on binge eating, and these effects were independent of age, BMI, the physical changes of puberty, and rGE processes. Findings were consistent across premenarcheal and postmenarcheal girls, and preliminary analyses confirmed that effects were due to estrogen, not progesterone or Estrogen \times\text{Progesterone} interactions, during the pubertal period. Overall, these results suggest that estrogen activation at puberty is associated with substantial etiologic shifts for binge eating that may have longer lasting effects on binge eating risk into adulthood.

The direction of effects for estrogen was opposite to what we observed in a small pilot study (Klump, Keel, et al., 2010). In the current sample, genetic influences were significantly stronger in twins with lower estradiol levels, whereas in the pilot study, genetic effects were stronger in twins with higher estradiol levels. This pattern of findings replicated across several different ways of analyzing the data (i.e., in the full sample of twins, in premenarcheal twins only, with and without covariates, using saliva samples on the assessment day only vs. all three assessment days, etc.), suggesting that the results were robust to sample composition, analytic approach, and study design decisions. Taken together, it seems that differences in results could be due to the much smaller sample sizes in the pilot study and that study’s inability to include discordant twin pairs in analyses. The current study took full advantage of a larger sample of twins and modeled hormone effects in concordant and discordant twins using continuous twin moderation models that simultaneously examined linear and non-linear differences in etiologic effects in all twins. Using this much larger sample and more powerful analytic techniques, we found evidence that lower estradiol levels are associated with stronger genetic effects (rather than the reverse) in the full sample of twins as well as in premenarcheal twins only.

Clearly, more replication studies are needed to definitively confirm our findings. If results replicate, it will be important to understand how/why genetic estimates for binge eating increase across the pubertal period when estradiol levels are increasing. It is possible that estrogen does not play any role in pubertal increases in genetic influences, and that the estrogen effects we observed reflect completely independent processes. In that case, it may be that other factors that come online and/or increase during puberty (e.g., leptin, kisspeptins, growth hormone; Saenger, 2003; Sanchez-Garrido & Tena-Sempere, 2013) are driving pubertal effects on genetic risk for binge eating rather than estrogen. Although these other factors should be examined in future research, we think that estrogen’s primary and central role in pubertal development and regulation of gene transcription within the CNS make this possibility less likely. Perhaps more importantly, increases in genetic effects during puberty/gonadarche occur in girls but not in boys (Klump et al., 2012), further suggesting that female-specific pubertal factors (like estrogen) rather than processes present in girls and boys (e.g., leptin, kisspeptins, and growth hormone) likely contribute to changes in genetic effects during puberty in girls.

If estrogen is involved in these pubertal effects, then our findings point to a potentially more complicated picture for estrogen than originally proposed (Klump, 2013; Klump et al., 2017; Klump, Keel, et al., 2010). Our results suggest that increasing genetic effects on binge eating during puberty may be driven by girls with lower levels of estradiol relative to other girls at similar ages and levels of development. Prior to the activation of the reproductive axis during puberty, all girls have relatively low levels of estradiol; however, once the reproductive axis is activated during puberty, girls begin to vary in their estradiol levels, and these individual differences appear to translate into stronger genetic influences on binge eating in girls with relatively lower (rather than higher) hormone levels. From an evolutionary perspective, these findings make intuitive sense, as all girls experience increases in estrogen during puberty that drive normative development and reproductive success. What our current findings suggest are that lower levels of estradiol during puberty may disrupt this developmental process and increase genetic effects on binge eating that (potentially) lead to the later development of binge eating in adulthood. Interestingly, although genetic effects were not examined, a recent animal study showed that complete removal of estrogen (via ovariectomy) before puberty in female rats resulted in substantially increased rates of binge eating in adulthood (Klump, Sinclair, Hildebrandt, Serfinski, & Sisk, 2016). These findings provide additional support for our theory that lower levels of estradiol during puberty may disrupt normative developmental processes and increase risk for binge eating.

Importantly, these data also imply that pubertal estrogen may be protective against binge eating. Estrogen is protective against binge eating in adulthood. Higher levels of estradiol are associated with low phenotypic rates of binge eating (Klump et al., 2013, 2014) and lower heritability (i.e., ~12%; Klump et al., 2015). By contrast, lower levels of estradiol are associated with higher phenotypic rates (Edler et al., 2007; Klump et al., 2013, 2014) and higher heritability (~40–50%; Klump et al., 2015, 2016) of the same binge-eating phenotypes. Studies in adulthood have used strong cross-sectional (Klump et al., 2016) and longitudinal (i.e., across the menstrual cycle; Klump et al., 2013, 2014, 2015) designs to establish these effects, and they directly replicate 40–50 years of animal work showing that estradiol is protective against excessive food intake in adulthood (Asarian & Geary, 2013; Klump et al., 2017). Taken together, data across puberty, adulthood, and animal and human studies suggest that estrogen may be protective against binge eating, and that lower levels of estradiol are associated with increased genetic effects and increased phenotypic rates of binge eating in adulthood.

Critically, there is one difference in the effects of estrogen during puberty versus adulthood—significant phenotypic estrogen/binge eating associations exist in adulthood, but no significant phenotypic associations have been observed before/during puberty in our current analyses (see Table 1) or past human (Klump, Keel, et al., 2010) and animal work (Klump et al., 2017). These differential phenotypic effects imply the presence of organizational and activational effects of estrogen on binge eating risk across development. The organizational/activational hypothesis of gonadal hormone action predicts that (a) gonadal hormones organize risk for binge eating during puberty through changes in gene transcription and resulting changes in brain structure/function; and (b) these changes organize the brain to respond to circulating levels of hormones in adulthood which activate and/or influence the expression of behavior (Sisk & Zehr, 2005). Importantly, when hormone organizational/activational effects are present, modest phenotypic associations between gonadal hormones and behavior are observed during puberty when organizational changes are taking place, whereas significant phenotypic associations are present in adulthood (Sisk & Zehr, 2005). This pattern of phenotypic associations
completely mimics the pattern observed for estrogen and binge eating in puberty/adulthood and suggests that estrogen’s actions during puberty might be organizational in nature, whereas its effects in adulthood are activational.

Indeed, based on current and prior data, it seems likely that relatively lower levels of estradiol may organize risk for binge eating during puberty via genomic changes within the CNS that lead to differential organization of neural circuitry and significant binge eating/estrogen associations in adulthood. These genomic changes may involve decreased expression/transcription of genes involved in normative development within key brain regions (e.g., brain reward pathways) or other, yet-to-be identified processes. Given the polygenic nature of binge eating, and the range of genetic/neural processes involved in pubertal brain development, the genomic changes likely involve a wide array of genes and neural networks that may code for mood/negative affect (e.g., serotonin genes; Barth, Villringer, & Sacher, 2015; Hildebrandt et al., 2010; Ostlund et al., 2003), ingestive behavior/obesity (e.g., FTO, TAS1R2, TAS1R3; Grimm & Steinle, 2011; Micali, Field, Treasure, & Evans, 2015; Tanofsky-Kraff et al., 2009), neuropeptides/hormones involved in satiety and food intake (e.g., cholecystokinin [CCK], ghrelin, leptin; de Krom et al., 2007; Grimm & Steinle, 2011; Monteleone, Tortorella, Castaldo, Di Filippo, & Maj, 2007), and/or reward system neurocircuitry (e.g., dopamine, opioids; Barth et al., 2015; Becker, 2009; Craft, 2008). To account for our findings, these systems must be regulated by estrogen and show pubertal changes in gene expression. Although hormonal regulation is known for some of these genes/systems (e.g., serotonin, dopamine, leptin, CCK; Barth et al., 2015; Becker, 2009; Brown & Clegg, 2012; Geary, 2001; Hildebrandt et al., 2010; Ostlund et al., 2003), more is needed in terms of identifying estrogen modulation during puberty and/or other developmental milestones. Nonetheless, the presence of hormone-mediated, pubertal sculpting of the CNS in animals (Ahmed et al., 2008; Barth et al., 2015; Nuñez, Sodhi, & Juraska, 2002; Primus & Kellogg, 1991; Spear, 2000; Zehr, Todd, Schulz, McCarthy, & Sisk, 2006) and genetically mediated associations between estradiol levels and changes in neural structures (i.e., less estradiol associated with more neural changes) in girls (Herting et al., 2014; Lenroot et al., 2009; Peper, Brouwer, et al., 2009; Peper, Schnack, et al., 2009) provides indirect support for key aspects of our organizational/activation theory and the candidate systems that may be involved.

Before ending, we should comment on findings regarding estrogen moderation of environmental effects. Similar to previous work, we found evidence for increasing nonshared environmental influences across estradiol levels (Klump, Keel, et al., 2010). The specific factors contributing to these environmental differences remain unclear, but they could reflect nongenomic, membrane-mediated actions of estrogen receptors (Santollo, Yao, Neal-Perry, & Eitgen, 2012) that become more prominent as levels of estradiol increase. Importantly, increases in nonshared environmental factors are unlikely to reflect differential environmental experiences for the twins based on age, BMI, or pubertal physical changes, as our analyses controlled for all of these factors. We found minimal differences in shared environmental effects across estradiol levels in the full twin sample or premenarcheal twins. This was somewhat surprising given twin studies showing significant differences in shared environmental factors across pubertal development (Klump, Culbert, O’Connor, et al., 2017) and estradiol levels (Klump et al., 2015, 2016; Klump, Keel, et al., 2010). It is possible that we were underpowered to detect differences in these effects, given some evidence for shared environmental influences in the premenarcheal twin sample (i.e., more modest MZ/DZ twin correlation differences - see Table 2) and known difficulties of identifying genetic and shared environmental influences in classical twin models (Martin, Eaves, Kearney, & Davies, 1978). Additional studies with larger sample sizes are needed to clarify whether estrogen moderates shared environmental influences on binge eating during puberty.

We should note a few additional limitations of our work. First, we used a continuous, self-report measure of binge eating rather than interview-based diagnoses and/or symptoms. Although self-reports might overestimate binge eating (Berg, Peterson, Frazier, & Crow, 2011), the percentage of our twins (3.1%) scoring above the binge eating scale cut-off for individuals with BN (von Ranson et al., 2005) is on par with what would be expected. Moreover, our measure performs well across the age range of pubertal development (Luo et al., 2016), and range of hormone values (see Supplemental Table S1 in the online supplemental material) examined, and it reliably distinguishes between girls who binge eat (see Methods) and/or have BN versus controls (von Ranson et al., 2005). Nonetheless, given discrepant results between current analyses and past pilot study results (Klump et al., 2010), additional replication studies are needed.

Second, we focused on binge eating rather than binge-related disorders. Sample sizes for disorders were too small for analysis, but binge eating appears to be dimensional rather than categorical in nature (Luo et al., 2016), and heritabilities of binge eating and binge-related disorders are similar (e.g., Bulik, Sullivan, & Kendler, 1998). Clearly, more studies of clinical disorders with interview-based measures are needed, but initial data suggest that results might extend to these disorders and measures as well.

Third, we used salivary hormone measures instead of serum. We focused on saliva in order to increase participation rates, ease data collection, and ensure that we were measuring active/unbound hormones (rather than the bound and unbound hormones in serum). Nonetheless, there is less published literature on the reliability/validity of salivary hormone values, and normative data for girls in our age range and during puberty are lacking. Normative data would be helpful for determining if our findings are driven by twins with non-normative/abnormally low levels of estradiol or are present across the full spectrum of estradiol values. Given our findings of linear (rather than nonlinear) differences in genetic effects, it seems unlikely that our results are accounted for by girls with abnormally low estradiol levels, but normative data could help confirm these impressions. In addition, more data on reliability/validity of salivary measures could confirm that our findings are not unduly influenced by measurement error, particularly at the lower levels of estradiol. It is important to note, however, that measurement error loads on the nonshared environment in twin models, and thus would result in increased nonshared environmental influences at lower estradiol levels (where presumably, error of measurement would be greater). This pattern of results is opposite to what we observed, where genetic effects were stronger, and nonshared environmental influences were weaker, at lower estradiol levels. This pattern of results was observed when modeling estradiol levels continuously as well as categorically in the full
sample of twins and premenarcheal twins only (see Tables 2 and 3, and Figure 2 and Supplemental Figure S3). Given these results, it seems unlikely that measurement error unduly influenced our results, but additional replications using salivary and serum measures of estradiol are needed to confirm our findings.

Finally, our data were cross-sectional rather than longitudinal in nature and focused on relatively discreet hormone measurements (i.e., 1–3 days of hormone values). It is therefore unclear whether our findings extend to within-subject/sample changes in genetic effects across multiple pubertal hormone assessments. We controlled for many variables that could account for differences in etiologic effects across our hormone levels (i.e., age, BMI, physical changes of puberty), and longitudinal and cross-sectional studies in adulthood (Klump et al., 2015, 2016) have confirmed that estrogen effects exist both within- and between-groups. Nonetheless, longitudinal studies that collect multiple hormone and behavioral measures across pubertal development are needed to confirm that findings reflect changes in estrogen across puberty. Ideally, these studies would capitalize on the existence of multiple animal models of binge eating (Klump et al., 2017) and conduct longitudinal assessments in both girls and female animals to directly test our theories of estrogen’s organizational/activational effects and substantially advance our understanding of ovarian hormone influences on binge eating and binge-related disorders in females.

References


example of the Australian Twin ADHD Project. Twin Research, 5, 324–328. http://dx.doi.org/10.1375/136905202320906309