A Behavior Genetic Investigation of Adolescent Motherhood and Offspring Mental Health Problems

K. Paige Harden, Stacy K. Lynch, Eric Turkheimer, Robert E. Emery, and Brian M. D’Onofrio
University of Virginia

Mary D. Waldron and Andrew C. Heath
Washington University—St. Louis

Wendy S. Slutske
University of Missouri—Columbia

Dixie J. Statham and Nicholas G. Martin
Queensland Institute of Medical Research

The present study examines the relations between adolescent motherhood and children’s behavior, substance use, and internalizing problems in a sample of 1,368 children of 712 female twins from Australia. Adolescent motherhood remained significantly associated with all mental health problems, even when using a quasi-experimental design capable of controlling for genetic and environmental confounds. However, the relation between adolescent motherhood and offspring behavior problems and substance use was partially confounded by family background variables that influence both generations. The results are consistent with a causal relation between adolescent motherhood and offspring mental health problems, and they highlight the usefulness of behavior genetic designs when examining putative environmental risks for the development of psychopathology. The generalizability of these results to the United States, which has a higher adolescent birth rate, is discussed.

Keywords: adolescent motherhood, children-of-twins, teenage pregnancy, internalizing, externalizing

The most obvious correlate of adolescent childbearing is privation: Teenage mothers are twice as likely to be impoverished when adults (Hoffman, Foster, & Furstenberg, 1993) and more likely to receive welfare (Moore et al., 1993), although the extent to which these adverse economic outcomes are a product of adolescents’ disadvantaged background, versus the pregnancy itself, is debated (Geronimus & Korenman, 1992; Lee & Gramotnev, 2006). In addition to the negative socioeconomic correlates of adolescent childbearing, the children of adolescent mothers are at elevated risk for diverse forms of psychopathology, including depressive and anxiety disorders (Hofferth, 1987; Moore, Morrison, & Greene, 1997), antisocial behavior, and other externalizing disorders (Jenkins, Shapka, & Sorenson, 2006; Levine, Pollack, & Comfort, 2001; Nagin, Pogarsky, & Farrington, 1997; Spieker, Larson, Lewis, Keller, & Gilchrist, 1999; Wakschlag et al., 2000). The disparity in adjustment between children of adolescent and adult mothers seems only to widen over children’s life spans, with the most dramatic disparities evident in adolescence and adulthood (Brooks-Gunn & Furstenberg, 1986; Furstenberg, Brooks-Gunn, & Morgan, 1987).

Consequently, several authors have concluded that a mother’s age at her first birth influences her child’s cognitive and psychosocial development and eventual adult adjustment. This has not remained a purely academic hypothesis: American organizations with disparate political ideologies, such as The Heritage Foundation (The Heritage Foundation, 2004) and Planned Parenthood (Planned Parenthood, 2006), publicly advocate that women should delay childbirth to improve their own and their children’s lives. The relation between adolescent motherhood and child adjustment, however, may not be as simple as commonly characterized. In the present article, we use a family design to consider how adolescent childbearing is related to child psychological adjustment, namely, the extent to which a teenage mother’s “background” variables—both environmental risks and genetic liabilities—account for her children’s higher risk for mental health problems.

Moreover, we use an Australian sample to extend the limited body of research documenting associations between adolescent motherhood and offspring adjustment beyond the United States.
Although the adolescent birthrate in the United States (54.4 per 1,000 girls aged 15–19) is almost triple that of Australia (19.8; Singh & Darroch, 2000)—almost the highest in the industrialized world—the limited extant research on the children of adolescent mothers in Australia suggests they are at risk for similar adverse psychological (Shaw, Lawlor, & Najman, 2006) and economic outcomes (Australian Institute of Health and Welfare, 2004). The higher birthrate among American adolescents is primarily due to less access to comprehensive sexual education (Weaver, Smith, & Kippax, 2005), ineffective contraceptive use (United Nations, 1991), and lower abortion rates (Singh & Darroch, 2000) compared with Australian adolescents, rather than more frequent or precocious sexual activity (Coley & Chase-Lansdale, 1998).

Disentangling Maternal Age From Maternal Family Background

There is an essential difficulty in resolving the extent to which adolescent childbirth is a cause of offspring psychopathology: Adolescent mothers differ from adult mothers in many respects other than age. Obviously, one cannot randomly assign children to be born to adolescent versus adult mothers. Moreover, many selection factors predict both age at first birth and children’s functioning and may account for the relation between adolescent motherhood and child psychopathology. For example, impoverished women are more likely to give birth as adolescents (Geronimus, 2004), and poverty is associated with increased risk for antisocial behavior in childhood and adolescence (Christ et al., 1990). In lieu of experimental control, researchers have largely attempted to measure and statistically control for potential environmental confounds such as socioeconomic status (Christ et al., 1990) or race (MacGregor, 2000). While such epidemiological investigations have advanced researchers’ understanding of the roles that certain aspects of an adolescent mother’s background environment play in her children’s poorer adjustment, the use of statistical controls is necessarily limited by researchers’ inability to measure perfectly every possibly relevant covariate.

Rutter and his colleagues (Rutter, 2005; Rutter, Pickles, Murray, & Eaves, 2001; Rutter & Silberg, 2002) outlined several strategy useful for this common methodological impasse. Primarily among them was quasiexperimentation, to “pull apart” (i.e., reduce or eliminate covariation between) confounded variables. Geronimus, Korenman, and Hillemeier (1994) first used a quasi-experimental family design to pull apart adolescent childbearing from environmental background: comparing the firstborn offspring of biological sisters. In a comparison of the children of two sisters, one who gives birth as a teenager and one who delays childbirth, any differences between the children cannot be attributed to environmental variables shared by sisters, including sociodemographic characteristics. This comparison, therefore, is a rigorous test of whether a woman’s set of environmental background influences, rather than her age at birth per se, are responsible for the apparent consequences of teenage pregnancy. The results indicated that heightened risk for negative behavioral problems in offspring were common to the children of socioeconomically disadvantaged mothers in general, not specific to teenage mothers. In a replication of Geronimus et al.’s (1994) study, Turley (2003) found that when environmental selection factors are controlled by comparing the children of biological sisters discordant for early childbearing, the effect for early motherhood on child adjustment is greatly attenuated, if not eliminated. Mothers who delay childbirth, but who share a similarly disadvantaged family background with teenage mothers, have children at equal risk for academic difficulties and behavior problems as the children of teenage mothers.

Comparing the children of biological sisters improves on standard epidemiological studies by controlling for unmeasured environmental variables that are shared by sisters; nevertheless, this approach does not control for environmental variables that may differ between siblings of different ages, nor does it fully control for genetic background variables (50% of genetic factors differ between biological siblings). A number of twin studies have demonstrated that genetic differences between individuals partly account for population variation in timing of first birth (Kirk et al., 2001; Kohler, Rodgers, & Christensen, 2002; Neiss, Rowe, & Rodgers, 2002) and family size (Fisher, 1930; but see Williams & Williams, 1974). In addition, there are genetic influences on other milestones of sexual development predictive of age at first birth (UDry & Cliquet, 1982), including age at menarche (Chern, Gatewood, & Anderson, 1980; Kaprio et al., 1995; Rowe, 2002) and age at initiation of sexual intercourse (Bailey, Kirk, Zhu, Dunne, & Martin, 2000; Doughty, 2000; Dunne et al., 1997; Lyons et al., 2004; Martin, Eaves, & Eyseenck, 1977; Miller et al., 1999; Rodgers, Rowe, & Buster, 1999). Previous research with the sample used in the present study found that genetic differences accounted for 33% of the variation in whether a woman gave birth as a teenager, whereas environmental influences shared by twins accounted for 27%, and environmental influences not shared by twins accounted for 40% (Waldron, 2004). These results are consistent with a larger body of research demonstrating that differences in family and other social environments partly reflect genetic differences between individuals.

Therefore, children not only experience the environmental disadvantages involved in having a teenage mother but also inherit the genetic factors that influence the likelihood of adolescent childbearing. In other words, their environmental experiences are correlated with their genetic liabilities, a phenomenon referred to in the behavior genetics literature as passive gene–environment correlation (or rGE; Plomin & Bergman, 1991; Rutter & Silberg, 2002). Moreover, genetic factors influence diverse forms of psychopathology, including internalizing problems (Haberstick, Schmitz, Young, & Hewitt, 2005), externalizing problems (Arsenault et al., 2003; Scourfield, Van der Bree, Martin, & McGuffin, 2004; Slutske et al., 1997), and alcohol and substance use problems (Heath et al., 1997; Heath & Martin, 1994). If the same genetic factors influence reproductive timing in women and risk for psychopathology in offspring (an example of pleiotropy, wherein a single set of genetic factors is related to multiple phenotypes), then the poorer adjustment seen in children of adolescent mothers may be mediated, to some extent, by common genetic liabilities. It should be noted that the origin of a psychosocial risk indicator is not necessarily the same as the mechanism of its effect, as made obvious by an analogy of smoking: The personality predictors of smoking initiation are completely independent of the mechanisms by which smoking causes lung cancer (Rutter, Silberg, & Simonoff, 1993). Genetic influence on adolescent motherhood, therefore, does not necessarily indicate that association between adolescent motherhood and offspring mental
health is not environmentally mediated; rather, genetic effects remain a possibility to be examined empirically.

Advantages of the Present Study: The Children-of-Twins Design

We improve on previous research on the relation between adolescent motherhood and offspring mental health problems by using a children-of-twins design, which has several important advantages. First, previous family designs (Geronimus et al., 1994; Turley, 2003) have studied offspring from early childhood to early adolescence; however, previous authors have suggested that the most potent effects of adolescent motherhood are evident later in offspring’s life spans (Brooks-Gunn & Furstenberg, 1986; Furstenberg et al., 1987). Our sample comprised adolescent and adult offspring aged 14–39 and thus was well suited for detecting these later effects.

Second, the children-of-twins design (D’Onofrio et al., 2003; Heath, Kendler, Eaves, & Markell, 1985; Nance & Corey, 1976) improves on the first-cousin comparison design used by Geronimus et al. (1994) and Turley (2003) because it compares the children of twin sisters, who share a common family environment and a common set of genes. In a comparison of a pair of identical (or monozygotic; MZ) twin mothers and their respective children, in which Twin A is younger at the birth of her first child than Twin B, and the firstborn child of Twin A demonstrates more mental health problems than the firstborn child of Twin B, the association cannot be attributed to any family background variables—either environmental or genetic in origin—that make twin mothers similar. Moreover, using both MZ and dizygotic (or fraternal; DZ) twins allows one to differentiate between environmental or genetic background variables. A significant within-twin pair association, although it remains confounded by any environmental influences that differ within twin pairs and are related to both adolescent childbearing and offspring psychopathology, is thus consistent with a causal relation.

Third, the majority of previous studies have been limited to only one child per mother (but see Turley, 2003). Sampling multiple children per twin parent permits comparisons between siblings: If the child born when his or her mother is a teenager demonstrates more mental health problems than his or her siblings born later in the mother’s life, then the association cannot be attributed to any environmental or genetic variables shared by siblings (Dick, Johnson, Viken, & Rose, 2000; Rodgers, Cleveland, van den Oord, & Rowe, 2000). Considered together, the family comparisons permitted by the children-of-twins design provide a rigorous test of whether a putative environmental risk—in this case, adolescent motherhood—causes adverse child outcomes. The children-of-twins design has been used previously to examine the processes by which marital conflict (Harden, Turkheimer, Emery, et al., 2007), stepfathering (Mendle et al., 2006), harsh punishment (Lynch et al., 2006), smoking during pregnancy (D’Onofrio et al., 2003), marital dissolution (D’Onofrio et al., 2005, 2006), parental schizophrenia (Gottesman & Bertelsen, 1989), and parental alcohol problems (Jacob et al., 2003) are related to offspring adjustment.

In the present investigation, we examine the relation between adolescent motherhood and behavior, substance use, and internalizing problems in adolescent and adult offspring. Although our investigation uses twins and their offspring, and thus may be broadly characterized as “behavior genetic,” we wish to emphasize that our goal was not to investigate genetic processes per se. Rather, our goal was to discriminate possible causal relations between adolescent motherhood and offspring mental health problems, similar to both previous epidemiological research using statistical controls (e.g., Christ et al., 1990) and to previous quasi-experimental family research (Geronimus et al., 1994; Turley, 2003). To that end, we capitalize on the clustering within families of potential confounding variables—both environmental and genetic in origin—by comparing biologically related children who differ in their exposure to adolescent motherhood.

Method

Participants

Participants included twins and their offspring from a volunteer twin register, formed in 1978 and maintained by the Australian National Health and Medical Research Council (NHMRC). Figure 1 illustrates the relationships among three Australian twin subsamples. The primary subsample to be used in the present research is the maternal subsample, but explaining the composition of this subsample necessitates describing all three.

![Figure 1. Relations among three Australia Twin Register subsamples.](image-url)
First, in 1993–1995, 5,889 twins (86% response rate) were interviewed by telephone as part of an ongoing investigation of the genetics of alcoholism (Genetics of Alcoholism Twin Sample; Heath et al., 1997). Mean ages at the 1993 interview were 42.7 years for men (range = 28–89 years) and 44.8 years for women (range = 27–90 years). In keeping with the low proportion of ethnic minorities in the non-Aboriginal Australian population, the twins were almost exclusively of European ancestry. The sample mirrors other population demographics as well. Previous analyses have found no effects of self-selection for marital status, religious affiliation, frequency of church attendance, personality traits, mental illness, and abnormal behavior (Heath et al., 1997; Slutske et al., 1997). The Genetics of Alcoholism Sample does overrepresent MZ twins, overrepresent twins born before 1930, and underrepresent twins with less than an 11th grade education (Baker, Treloar, Reynolds, Heath, & Martin, 1996; Slutske et al., 1997). Cooperation bias and the underrepresentation of poorly educated participants, however, have been shown to have negligible effects on behavioral genetic analyses of conduct disorder (Heath et al., 1996). Overall, the Genetics of Alcoholism Sample can be considered broadly generalizable to the non-Aboriginal Australian population.

Between 1998 and 2001, investigators contacted a selected subsample of the 1993 interview sample’s offspring (children-of-twins sample). Selection targeted offspring whose parents reported in previous interviews histories of conduct disorder, depressive disorder, alcohol dependence, and/or divorce, as well as a control group whose parents reported none of the above. In total, 2,554 offspring of 1,409 adult twins participated in the study (an 82% response rate). Of the total sample of children-of-twins, 51% came from nuclear families in which the twin parent did not have a history of psychopathology or divorce; of these children, 47% (24% of the total sample) came from nuclear families in which neither the twin parent nor the cotwin had a history of psychopathology or divorce.

The purpose of the present study was to examine early motherhood, thus the maternal subsample was restricted to complete and incomplete female–female twin pairs and their offspring. This subsample consisted of 1,368 children (51.5% girls, 48.5% boys) of 230 complete pairs of twins and 252 individual twins. Women from incomplete twin pairs and their children were included in all analyses, unless noted otherwise, because the data were used for sibling comparisons. Of the twins, 60.1% were from MZ pairs and 39.04% from DZ pairs. The number of children per nuclear family ranged from 1 to 6; the mean number of siblings per nuclear family was approximately 2. The offspring’s age at assessment ranged from 14 to 39 years ($M = 25.1, SD = 5.7$). The maternal subsample was used in the present analyses.

**Measures**

Zygosity was determined by questionnaire responses concerning physical similarity and frequency of occasions in which twins were mistaken for each other. When there was disagreement between cotwins about zygosity or when zygosity assignment was otherwise ambiguous, further information, including photographs, was requested. Comparisons of these zygosity assignments with multilocus genotyping have shown the self-report questions to be greater than 95% accurate (Eaves, Eysenck, & Martin, 1989). In addition, final zygosity assignments from questionnaire responses demonstrated perfect agreement with zygosity assignment on the basis of DNA typing of eight polymorphic markers in a subsample of 190 twin pairs (Duffy, 1994).

Mothers’ age at each birth and children’s psychosocial adjustment were assessed using the Semi-Structured Assessment for the Genetics of Alcoholism–OZ version (SSAGA-OZ; Bucholz et al., 1994), a comprehensive psychiatric interview designed for genetic studies of alcoholism, modified for use over the telephone. The SSAGA-OZ is derived from the National Institute of Mental Health Diagnostic Interview Schedule (DIS; Helzer & Robins, 1988), the Structured Clinical Interview for DSM–III–R (Spitzer, Williams, Gibbon, & First, 1992), the Schedule for Affective Disorders and Schizophrenia (SADS; Endicott & Spitzer, 1978), the Composite International Diagnostic Interview (Robins et al., 1988), and the HELPER Interview (Coryell, Cloninger, & Reich, 1978). The SSAGA-OZ includes Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM–IV; American Psychiatric Association, 1994) diagnostic items for oppositional defiant disorder, attention-deficit/hyperactivity disorder, conduct disorder, alcohol dependence and abuse, illicit drug use, and major depression. Additional items assess suicidal intent and self-injurious behavior, lifetime cigarette use, regular smoking, lifetime history of alcohol use, and regular alcohol use. Interrater reliability of the SSAGA-OZ has been shown to be excellent and did not appear to be compromised by telephone use (Bucholz et al., 1994; Heath et al., 1997). Interviews were administered by trained lay interviewers who were supervised by a trained clinical psychologist. All interviews were audiotaped, and a randomly selected 5% of the interviews were reviewed for quality control and check of coding inconsistencies.

**Correction for Sample Selection**

Similar to many genetic epidemiological studies, the maternal subsample is a product of deliberate selection on stratification variables (parental psychiatric history and marital status), in conjunction with potentially nonrandom self-selection. Without addressing this sample selection problem, our analyses may be biased. In general, sample selection may be considered a case of missing data, with data on the variables of interest only present in selected twins and their families (Bechger, Boomsma, & Koning, 2002). Accordingly, sample selection may be considered within Rubin’s model of missing data (Little & Rubin, 1987; Rubin, 1976). Missingness (i.e., selection) is considered ignorable not only when participants are a random sample from the general population (selected completely at random SCAR) but also when selection depends on the values of other variables, related to the variables of interest, observed in both selected and unselected participants (selected at random; SAR).

As detailed above, the maternal subsample is derived from a larger, population-representative twin sample (the Genetics of Alcoholism Twin Sample), in which multiple sociodemographic and psychiatric characteristics, including the deliberate selection variables, were observed for both selected and unselected twins. Therefore, the maternal subsample may be considered SAR. Following Heath, Madden, and Martin’s (1998) procedure for devel-
opling and testing models of nonresponse in SAR data, we used multiple logistic regression to identify predictors of whether a female–female twin pair (i.e., at least one twin) from the Genetics of Alcoholism Twin Sample participated in the maternal subsample. Pairwise participation, rather than individual twin participation, was predicted because sample selection occurred on the pair level. Sociodemographic and psychiatric characteristics assessed using the SSAGA-OZ were used as predictors of selection into the maternal subsample (see Bucholz et al., 1994; or Heath et al., 1998, for more details). Propensity weights were then constructed as the inverse predicted probability of a pair participating in the maternal subsample, as calculated from the logistic regression model (for details on standard methods of data weight construction, see Heath et al., 1998; or Lee, Fonthofor, & Lorimor, 1989). Our model for selection was tested by comparing the unweighted and weighted frequency distributions of sociodemographic and psychiatric characteristics in the maternal subsample with the distributions in the female–female twin pairs from the Genetics of Alcoholism Sample.

Table 1 summarizes the observed frequency distributions for sociodemographic characteristics, family history of alcohol use and depression, alcohol use behaviors, and psychiatric disorders of twin pairs participating in the maternal subsample compared with nonselected female–female twin pairs from the population-representative Genetics of Alcoholism Sample. On the basis of univariate comparisons, selected twin pairs were different from nonselected female pairs in multiple respects, including birth cohort, family history of alcohol and depression, panic, phobia, high-frequency drinking, suicidal ideation, major depression, alcohol dependence, conduct disorder, and migraine. Also summarized in Table 1 are partial odds ratios estimated from a multiple logistic regression predicting participation in the maternal subsample. The following were significant predictors of sample selection: birth cohort, age, discordance for spouse alcohol use history, twin and spouse depression history, panic, agoraphobia, social phobia, major depressive disorder, alcohol dependence, conduct disorder, concordance for paternal alcohol use history, high maximum drinking, and migraine and major depressive disorder. Most of these predictors were deliberate selection variables.

Finally, the right-hand columns of Table 1 summarize the frequency distributions for the maternal subsample when pairwise propensity weights were used. A comparison of these distributions may be used to examine whether data weighting removed differences between selected and nonselected participants. Of the significant predictors of sample selection identified by multiple logistic regression, differences in mean age, as well as frequency differences in birth cohort, discordance for panic, social phobia, major depressive disorder, and conduct disorder, and concordance for high maximum number drinks, were removed by data weighting. In addition, frequency differences in discordance for spouse alcohol use history, twin and spouse depression history, and alcohol dependence, and concordance for migraine, major depression, and paternal alcohol use history were substantially reduced, although univariate comparisons remained statistically significant. Residual differences in the frequency of psychiatric disorders may be evident even after data weighting because the probability of selection was heavily driven by sociodemographic characteristics, partic-

ularly birth cohort, as illustrated by the extreme odds ratios associated with birth cohort dummy variables. These propensity weights were used in all subsequent analyses. Despite data clustering, all methods for single-level data weighting are applicable because weights are applied at the highest level of clustering only (Asparouhov, 2004).

Exploratory Factor Analysis (EFA) of Offspring Adjustment

We performed an EFA of dichotomous DSM–IV diagnostic symptoms endorsed by children from the entire children-of-twins sample during the SSAGA-OZ interview. Only one child per twin pair from the entire children-of-twins sample was included to avoid bias introduced by response correlations between related participants. One SSAGA item (forcing another into sexual intercourse) was dropped because of low endorsement frequency. All EFAs were conducted in Mplus (Muthén & Muthén, 1998–2004) using Promax rotation. The EFA resulted in a three-factor solution, with interfactor correlations between .39 and .49. Each factor showed high internal consistency (Factor 1 $\alpha = .90$, Factor 2 $\alpha = .87$, Factor 3 $\alpha = .91$).¹

The first factor, Behavior Problems (BP), includes reports of oppositional defiant behaviors, attention problems, hyperactivity, conduct problems (excluding serious violations of rules), and report of recurrent legal problems due to alcohol use. The second factor, Substance Use Problems (SUP), includes cigarette use, alcohol use, alcohol abuse, alcohol dependence, drug use, problems associated with drug use, and serious violations of rules. The third factor, Internalizing Problems (IP), includes depressive episode criteria and items concerning suicide. A 1-year follow-up of 176 of the offspring of the twins found the child adjustment factors to be reliable ($\kappa = 0.78$ for BP, $0.89$ for SUP, and $0.75$ for IP). We summed the endorsed items loading on each factor to calculate unit-weighted factor scores for children’s BP ($M = 3.52$, $SD = 4.42$, range $= 0–28$), SUP ($M = 10.26$, $SD = 6.02$, range $= 0–32$), and IP ($M = 4.03$, $SD = 4.39$, range $= 0–16$). Symptom counts were used rather than clinical diagnoses because clinical-level pathology was rare; symptom counts better capture the range of outcomes expected in the general population.

Results

Descriptive Means Comparisons

All descriptive means comparisons were conducted using the propensity weights to correct for sample selection bias. We conducted five descriptive means comparisons to illustrate the between-family, within-twin pair, and within-nuclear family associations. These comparisons of means were meant to be purely illustrative, thus no inferential statistics were computed. The role of sampling error is assessed in the following, more rigorous, hierarchical linear modeling analyses. In lieu of probability testing,

¹ Complete details of the EFA are available upon request from the first author or from D’Onofrio et al. (2005).
### Table 1
Comparison of Sociodemographic and Psychiatric Characteristics of Female Twins From the Genetics of Alcoholism Sample and the Maternal Subsample, Before and After Data Weighting

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Genetics of alcoholism sample (%)</th>
<th>Maternal subsample (%)</th>
<th>Odds ratios&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Weighted maternal subsample (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sociodemographic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Under 35 years</td>
<td>26.80</td>
<td>9.54&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.0</td>
<td>27.59</td>
</tr>
<tr>
<td>35–45 years</td>
<td>32.47</td>
<td>45.64&lt;sup&gt;a&lt;/sup&gt;</td>
<td>15.16 (9.84–23.36)</td>
<td>26.51</td>
</tr>
<tr>
<td>Over 45 years</td>
<td>40.73</td>
<td>44.81&lt;sup&gt;a&lt;/sup&gt;</td>
<td>52.45 (27.08–101.59)</td>
<td>35.89</td>
</tr>
<tr>
<td>Age (Mean)</td>
<td>45.25</td>
<td>44.92</td>
<td>0.90 (0.88–0.93)</td>
<td>43.26</td>
</tr>
<tr>
<td><strong>Discordant pairs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family alc. use history</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twin</td>
<td>7.73</td>
<td>8.09</td>
<td>—</td>
<td>10.22</td>
</tr>
<tr>
<td>Father</td>
<td>10.30</td>
<td>9.75</td>
<td>—</td>
<td>12.24</td>
</tr>
<tr>
<td>Mother</td>
<td>2.68</td>
<td>2.49</td>
<td>—</td>
<td>2.25</td>
</tr>
<tr>
<td>Spouse</td>
<td>17.81</td>
<td>25.52&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.71 (1.24–2.34)</td>
<td>19.17&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Family depr. history</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twin</td>
<td>32.93</td>
<td>40.46&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.35 (1.01–1.80)</td>
<td>35.56&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Father</td>
<td>16.66</td>
<td>17.43</td>
<td>—</td>
<td>16.93</td>
</tr>
<tr>
<td>Mother</td>
<td>22.41</td>
<td>24.07</td>
<td>—</td>
<td>21.20</td>
</tr>
<tr>
<td>Spouse</td>
<td>22.79</td>
<td>31.95&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.59 (1.20–2.11)</td>
<td>24.95&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Agoraphobia</td>
<td>6.52</td>
<td>6.64</td>
<td>0.45 (0.26–0.78)</td>
<td>6.11</td>
</tr>
<tr>
<td>Panic</td>
<td>6.85</td>
<td>11.62&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.79 (1.10–2.92)</td>
<td>6.96</td>
</tr>
<tr>
<td>Social phobia</td>
<td>3.62</td>
<td>5.39&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—</td>
<td>4.16</td>
</tr>
<tr>
<td>Other phobias</td>
<td>2.85</td>
<td>4.56&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—</td>
<td>3.55</td>
</tr>
<tr>
<td>Abstain from alcohol</td>
<td>4.11</td>
<td>3.73</td>
<td>—</td>
<td>4.19</td>
</tr>
<tr>
<td>High-frequency drinking</td>
<td>23.62</td>
<td>27.18&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—</td>
<td>23.21</td>
</tr>
<tr>
<td>High-density drinking</td>
<td>12.82</td>
<td>12.03</td>
<td>—</td>
<td>13.87</td>
</tr>
<tr>
<td>High-maximum number drinks</td>
<td>18.14</td>
<td>17.22</td>
<td>—</td>
<td>17.12</td>
</tr>
<tr>
<td>Migraine</td>
<td>38.58</td>
<td>35.48</td>
<td>—</td>
<td>39.04</td>
</tr>
<tr>
<td>Suicidal ideation</td>
<td>4.52</td>
<td>7.26&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—</td>
<td>5.99</td>
</tr>
<tr>
<td>Major depression</td>
<td>33.59</td>
<td>43.36&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.05 (1.53–2.76)</td>
<td>34.18</td>
</tr>
<tr>
<td>Alcohol dependence</td>
<td>8.60</td>
<td>12.24&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.92 (1.20–3.08)</td>
<td>10.64&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Conduct disorder</td>
<td>3.34</td>
<td>5.81&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2.92 (1.47–5.80)</td>
<td>3.24</td>
</tr>
<tr>
<td><strong>Concordant pairs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family alc. use history</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twin</td>
<td>0.38</td>
<td>1.04&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—</td>
<td>0.43</td>
</tr>
<tr>
<td>Father</td>
<td>10.36</td>
<td>16.80&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.76 (1.19–2.59)</td>
<td>12.82&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Mother</td>
<td>2.08</td>
<td>3.32&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—</td>
<td>2.60</td>
</tr>
<tr>
<td>Spouse</td>
<td>1.75</td>
<td>2.70</td>
<td>—</td>
<td>1.53</td>
</tr>
<tr>
<td>Family depr. history</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twin</td>
<td>9.86</td>
<td>17.22&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—</td>
<td>11.54&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Father</td>
<td>6.90</td>
<td>10.37&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—</td>
<td>8.61&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Mother</td>
<td>11.89</td>
<td>19.09&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—</td>
<td>12.46</td>
</tr>
<tr>
<td>Spouse</td>
<td>2.47</td>
<td>4.36&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—</td>
<td>2.44</td>
</tr>
<tr>
<td>Agoraphobia</td>
<td>0.71</td>
<td>2.07&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—</td>
<td>0.89</td>
</tr>
<tr>
<td>Panic</td>
<td>0.55</td>
<td>0.62</td>
<td>—</td>
<td>0.37</td>
</tr>
<tr>
<td>Social phobia</td>
<td>0.11</td>
<td>0.21</td>
<td>—</td>
<td>0.08</td>
</tr>
<tr>
<td>Other phobias</td>
<td>0.11</td>
<td>0.21</td>
<td>—</td>
<td>0.08</td>
</tr>
<tr>
<td>Abstain from alcohol</td>
<td>0.88</td>
<td>1.04</td>
<td>—</td>
<td>0.82</td>
</tr>
<tr>
<td>High-frequency drinking</td>
<td>8.71</td>
<td>8.92</td>
<td>—</td>
<td>11.00</td>
</tr>
<tr>
<td>High-density drinking</td>
<td>1.97</td>
<td>2.49</td>
<td>—</td>
<td>1.70</td>
</tr>
<tr>
<td>High-maximum number drinks</td>
<td>4.88</td>
<td>4.36</td>
<td>0.47 (0.22–0.995)</td>
<td>6.72</td>
</tr>
<tr>
<td>Migraine</td>
<td>28.22</td>
<td>40.46&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.59 (1.16–2.19)</td>
<td>32.58&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Suicidal ideation</td>
<td>0.44</td>
<td>1.04&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—</td>
<td>0.53</td>
</tr>
<tr>
<td>Major depression</td>
<td>11.51</td>
<td>19.92&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.83 (1.15–2.92)</td>
<td>16.02&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Alcohol dependence</td>
<td>1.10</td>
<td>1.87</td>
<td>—</td>
<td>1.88</td>
</tr>
<tr>
<td>Conduct disorder</td>
<td>0.44</td>
<td>0.83</td>
<td>—</td>
<td>0.52</td>
</tr>
</tbody>
</table>

**Note.** Dashes represent odds ratios that are not significantly different from zero (that is, not applicable). alc. = alcohol; depr. = depression.

<sup>a</sup> Univariate comparisons with nonselected female participants from the Genetics of Alcoholism Sample are significant at *p < .05.* <sup>b</sup> All listed odds ratios are significantly different from 1; *p < .05. 95% confidence intervals are shown in parentheses. Nonsignificant odds ratios for remaining predictors are not shown.
we computed effect sizes \( d \) (Cohen, 1988) for each comparison.\(^2\) Teen mothers were defined as women who were \( \leq 20 \) years old at childbirth. The five comparisons were as follows:

1. All children born to teen mothers \((n = 91)\) versus all children born to nonteen mothers \((n = 1,277)\).
2. All children born to teen mothers versus firstborn children of nonteen mothers \((n = 471)\).
3. All children born to teen mothers versus later born (i.e., not firstborn) children born to nonteen mothers who had their first child when still a teen \((n = 329)\).
4. Laterborn siblings of children born to teen mothers versus laterborn children born to nonteen mothers who never had a teen birth \((n = 477)\).
5. Firstborn children classified into four groups: (a) children of nonteen mothers whose aunts never had a teen birth \((n = 238)\); (b) children of nonteen mothers whose aunts had a teen birth \((n = 28)\); (c) children of teen mothers whose aunts never had a teen birth \((n = 28)\); and (d) children of teen mothers whose aunts also gave birth as a teen \((n = 14)\). (Only firstborn children of complete twin pairs were included because information on both twins is necessary to accurately classify children into groups.)

Table 2 shows the mean BP, SUP, and IP of children grouped by the age of their mother. Comparing the first two columns on the left of Table 2, the children born to mothers \( \leq 20 \) years old demonstrated more BP \((d = 0.435)\), SUP \((d = 0.414)\), and IP \((d = 0.251)\) than children of older mothers. Comparing the first and third columns of Table 2, the children born to teenage mothers demonstrated higher BP \((d = 0.535)\), SUP \((d = 0.513)\), and IP \((d = 0.276)\) than the firstborn children of older mothers. Of the 91 children born to teen mothers, the majority were firstborn children \((n = 76)\), and all but 1 were second born at the latest. This comparison, therefore, partly eliminated the effects of birth order. These phenotypic associations, the latter partially controlling for birth order effects, are consistent with previous epidemiological research.

The third analysis compares the children of teenage mothers with their later born siblings, thus controlling for any genetic or environmental variables common to children in the same nuclear family. Comparing the first and fourth columns of Table 2, children born to teenage mothers show more BP \((d = 0.334)\), SUP \((d = 0.292)\), and IP \((d = 0.177)\) than siblings born later in their mother’s development. This is consistent with a causal effect, although the effect sizes are attenuated compared with the between-family comparisons. The relation between adolescent motherhood and child mental health problems may be attenuated when comparing siblings rather than unrelated children because some of the epidemiological association is due to environmental or genetic variables common to children in a nuclear family, or because adolescent childbearing changes the nuclear family environment such that all subsequent children are also adversely affected (e.g., by reducing later socioeconomic resources). Caution in interpreting the sibling comparison is warranted.

The fourth analysis compares children who were all born to nonteen mothers \((M_{age} = 28.3 \text{ years})\) in the first group vs. \(M_{age} = 28.9 \text{ years}\) in the second group) and who have similar birth orders (all nonfirstborn), but who differ in whether they were born to a mother who had her first child as a teen and thus have different levels of environmental and genetic risk. In other words, they only differ in whether they come from the type of nuclear family that is started when a woman is still a teenager. Comparing the fourth and fifth columns of Table 2, later born children whose mother had her first child when still a teen did not demonstrate substantially more BP \((d = -0.028)\), SUP \((d = 0.061)\), or IP \((d = 0.072)\). This, too, is consistent with a causal effect because an increased risk of psychopathology was not evident for offspring not born when their mothers were teenagers.

The final means analysis (see Table 3) compared children divided into groups on the basis of their aunt’s and their mother’s teen birth status. Because information on both twins was necessary to classify children into groups, only children from complete twin pairs were included in the final means analysis. The key comparison is between the children of discordant twin pairs (i.e., between first cousins in discordant twin families): Does the twin who gave birth earlier than her cotwin have a child who demonstrates poorer adjustment than his or her cousin? In other words, does being born into the type of extended family in which women give birth as teenagers, even if one’s own mother delays childbearing, predict one’s adjustment as well as being born to a teen mother? Com-

\[
\bar{x}_2 - \bar{x}_1 \pm t_{\alpha, n-1} \frac{s_p}{\sqrt{n_1 + n_2}}
\]

An effect size of 0.2 is considered small, 0.5 medium, and 0.8 large.
paring the second and third columns of Table 3, children born to teenage mothers demonstrated substantially more BP (\(d = 0.214\)), SUP (\(d = 0.408\)), and IP (\(d = 0.383\)) than their first cousins born to older mothers. This pattern of results, again, is consistent with causal effects of early motherhood, rather than the operation of genetic or environmental third variables.

Hierarchical Linear Models

The above means comparisons, although capitalizing on multiple family relationships, have three critical limitations. First, the inclusion of varying numbers of nonindependent observations per nuclear family weights some families more heavily than others, precluding the correct estimation of standard errors. Second, the means comparisons do not capitalize on the difference in genetic relatedness between MZ and DZ twin pairs and are thus incapable of distinguishing between environmental and genetic confounds. Third, means comparisons, by definition, illustrate the relations between adolescent childbearing and mean numbers of behavior, substance use, and internalizing problems, but they do not illustrate any relations with the extent to which children in the same family differ in mental health problems. HLM is an analytic strategy capable of addressing the above limitations in two ways (see Raudenbush & Byrk, 2002, for a review). First, HLM explicitly models nonindependence among observations by partitioning unaccounted for variance into that shared by members of a cluster \(\tau_{ijk}\), variation shared by all children within a nuclear family (i.e., siblings; \(\tau_{0jk}\)), and residual variation unique to individual offspring (\(\sigma^2\)). (Dividing the variation shared by a cluster [e.g., \(\tau_{0jk}\)] by the total variance yields the familiar intraclass correlation.)

Similar to traditional regression analysis, this model can be first expanded to include the effect of being born to a teenage mother and subsequently expanded to include measured covariates as statistical controls:

\[
y_{ijk} = B_{000} + B_{1}x_{ijk} + \ldots + B_{p}x_{ijk} + \tau_{ijk} + \epsilon_{ijk}. \tag{2}
\]

In the above model, \(x_{ijk}\) is a dichotomous variable representing whether a child was born to a teenage mother; other \(x\) variables include the effect of child gender and the linear and quadratic effects of child age.

The methodological controls afforded by the children-of-twins design may be modeled by including additional fixed effects. First, we may include a dichotomous variable representing whether a mother’s first child was born when she was a teenager (\(X_{0ijk}\)):

\[
y_{ijk} = B_{000} + B_{1}x_{1ijk} + B_{2}x_{0ijk} + \ldots + B_{p}x_{pijk} + \tau_{00k} + \tau_{ijk} + \epsilon_{ijk}. \tag{3}
\]

As shown in Figure 2, the complex organization of our sample can be considered on three levels: Individual offspring comprise Level 1; clusters of offspring born to an individual mother (siblings in nuclear families) comprise Level 2; and clusters of offspring born to both mothers in a twin pair (cousins in twin families) comprise Level 3. The most basic HLM for our three-level data is an unconditional model:

\[
y_{ijk} = B_{000} + \tau_{00k} + \tau_{ijk} + \epsilon_{ijk}, \tag{1}
\]

\[
\tau_{00k} \sim N(0, \tau_{000}); \tau_{ijk} \sim N(0, \tau_{111}); \epsilon_{ijk} \sim N(0, \sigma^2).
\]

which models the adjustment of the \(i\)th child of the \(j\)th twin in the \(k\)th twin pair as a function of an overall intercept, variation shared by all children within a twin family (i.e., cousins; \(\tau_{00k}\)), variation shared by all children within a nuclear family (i.e., siblings; \(\tau_{ijk}\)), and residual variation unique to individual offspring (\(\sigma^2\)).

Table 3

<table>
<thead>
<tr>
<th>Problem</th>
<th>No teen mother/no teen aunt</th>
<th>No teen mother/teen aunt</th>
<th>Teen mother/no teen aunt</th>
<th>Teen mother/teen aunt</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(M)</td>
<td>(SE)</td>
<td>(M)</td>
<td>(SE)</td>
</tr>
<tr>
<td>Behavior</td>
<td>3.00</td>
<td>0.28</td>
<td>2.84</td>
<td>0.73</td>
</tr>
<tr>
<td>Substance use</td>
<td>9.60</td>
<td>0.37</td>
<td>8.95</td>
<td>1.19</td>
</tr>
<tr>
<td>Internalizing</td>
<td>3.81</td>
<td>0.25</td>
<td>5.26</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Note. Italic models nonindependence among observations by partitioning unaccounted for variance into that shared by members of a cluster \(\tau_{ijk}\), variation shared by all children within a nuclear family (i.e., siblings; \(\tau_{00k}\)), and residual variation unique to individual offspring (\(\sigma^2\)).
Notice that the inclusion of this Level 2 covariate changes the interpretation of the $x_{ijk}$ effect: Among children born to mothers who had their first child when still a teenager, does being born oneself to a teenage mother predict adjustment? This model more precisely recapitulates the descriptive means analysis comparing children of teenage mothers with their later born siblings. Second, we may include an ordinal variable ($x_{00k}$) representing how many women in a twin family gave birth to their first child while still a teen (0, 1, or 2):

$$y_{ijk} = B_{000} + B_{1}x_{1ijk} + B_{2}x_{0ijk} + B_{3}x_{00k}$$

$$+ \ldots + B_{n}x_{nkijk} + r_{ijk} + r_{ijk} + e_{ijk}. \quad (4)$$

Again, the inclusion of this Level 3 covariate changes the interpretation of the $x_{00k}$ effect: Among children born into the type of extended family in which women are at risk for teenage pregnancies, does being born to mother who gives birth as a teen predict adjustment? This model more precisely recapitulates the final descriptive means analysis comparing the firstborn children of twins discordant for teenage childbearing.

Third, a significant effect of the Level 3 covariate ($x_{00k}$) indicates that being born into an extended family in which women give birth as teenagers, regardless of whether one’s own mother gave birth as a teenager, predicts poorer adjustment. Whether this effect is due to environmental or genetic variables can be estimated by including an interaction between the twin family covariate and zygoty (coded as 0 in MZ twin families and 1 in DZ twin families):

$$y_{ijk} = B_{000} + B_{1}x_{1ijk} + B_{2}x_{0ijk} + B_{3}x_{00k} + B_{4}z_{ygo} + B_{5}z_{ygo}$$

$$+ \ldots + B_{n}x_{nkijk} + r_{ijk} + r_{ijk} + e_{ijk}. \quad (5)$$

If, for example, the relationship between early motherhood and BP were entirely due to transmission of common genetic liabilities from parent to child, then children’s BP will be predicted equally well by their aunt’s teenage childbearing as their mother’s in MZ twin families. In DZ families, however, children’s BP will be better predicted by their mother’s teenaged childbearing because of genetic differences between sisters. By coding zygoty with 0 for MZ and 1 for DZ, the regression coefficient for $x_{00k}$ ($B_3$) represents the relation between number of teen births in the twin family and offspring mental health problems in MZ twin families, and the regression coefficient for the interaction term ($B_4$) represents the difference in the effect of the twin family covariate between MZ and DZ twin families. If the influences confounding the associations between early motherhood and adjustment were genetic in origin, then we would expect the effect of the twin family covariate to be twice as large in MZ families as DZ families because MZ twins share twice as many genes as DZ twins. (It is conventional to include the main effect of an interaction variable, but we do not expect overall adjustment to differ significantly between the children of MZ and DZ twin families.) All models were estimated in SAS using the propensity weights to correct for sample selection bias. Data from complete and incomplete twin pairs and their children were used. Model fit was compared using differences in model chi-square, which are themselves distributed as chi-square, with degree of freedom equal to the difference between the models’ degree of freedom.

### Behavior Problems

Results from the unconditional model (not tabulated) indicate that cousins are not reliably similar in their BP ($\tau_{000} = 0$). This is consistent with previous analyses of children-of-twin analyses using this data set (e.g., Lynch et al., 2006). The intraclass correlation reflects the extent to which siblings are similar; it is calculated as the proportion of unaccounted variance shared by members of a cluster. The nuclear family intraclass correlation for BP was 0.189, indicating that there was only a small degree of similarity among siblings.

Model 1, shown in Table 4, showed there was a significant effect of being born to a teenage mother. Model 2 statistically controlled for the effect of child gender and the linear and quadratic effects of child age. Boys reported significantly more BP, consistent with previous research; however, there were no significant effects of age. Even after the addition of these covariates, being born to a teenage mother remained a significant predictor of BP. Model 3 included the effect of being born to a mother who had her first child as a teen; this covariate had no significant effect. The effect of being born to a teen mother remained significant, consistent with a causal hypothesis. Model 4 included the effect of the number of teen births in the twin family. This covariate significantly predicted BP, indicating that environmental or genetic variables shared by twin sisters were responsible for at least part of the association between early motherhood and BP. Nevertheless, the effect of being born to a teen mother remained significant, again consistent with a causal hypothesis.

Next, whether the family level confounds were environmental or genetic in origin was tested with an interaction between the number of teen births in the teen family and the zygosity of the teen pair in Model 5. The direction and magnitude of the interaction was consistent with the hypothesis that genetic transmission accounts, to some extent, for the relation between adolescent motherhood and offspring mental health problems: The effect of the twin family covariate was approximately twice as large in MZ twin families as DZ families. The interaction term was marginally nonsignificant ($p = .06$), but the change in model fit from Model 4 to Model 5 was significant, thus it was somewhat ambiguous whether the difference between MZ and DZ families could be attributed to sampling error. Overall, the results suggest that BP remains significantly related to adolescent motherhood, even when comparing family members with control for genetic and environmental confounds. Figure 3 illustrates how the effect of being born to a teenage mother on BP was attenuated—but did not disappear—when comparing siblings and first cousins, rather than unrelated children.

### Substance Use Problems

Results from the unconditional model (not tabulated) indicated that, unlike behavior problems, cousins were modestly similar in their SUP (intraclass correlation = 0.195; $\tau_{000} = 6.29$). Siblings were even more similar in their SUP (intraclass correlation = 0.249; $\tau_{11} = 6.45$). Subsequent HLM, shown in Table 5, revealed a pattern similar to that seen for behavior problems. Male children reported more SUP. Older children reported more SUP, with the effect of age tapering as offspring reached middle age. Regardless of the statistical or methodological controls added, the effect of being born to a teenage mother remained significant. In addition, the number of teen births in the twin family predicted SUP. The
magnitude and direction of the interaction term (approximately twice as large in MZ families as DZ families) was consistent with genetic selection effects. However, the change in model fit from Model 4 to Model 5 was marginally nonsignificant (\(p = 0.06\)), as was the coefficient itself, suggesting that the analyses lacked the power to discriminate between environmental and genetic effects.

Similar to Figure 3, Figure 4 illustrates how the effect of being born to a teenage mother decreases when comparing siblings and first cousins, rather than unrelated children.

Internalizing Problems

Results from the unconditional model (not tabulated) indicated that there was little reliable similarity between cousins (intraclass correlation = 0.065; \(\tau_{000} = 1.04\)) and between siblings (intraclass correlation = 0.183; \(\tau_{111} = 2.94\)). Most of the variation in IP, then, exists among siblings. Subsequent HLM, shown in Table 6, revealed a pattern somewhat different than behavior problems and substance use. Male children reported fewer IPs, consistent with previous research. There were no effects of age. Regardless of the statistical or methodological controls added, the effect of being born to a teenage mother remained significant, consistent with the causal hypothesis. Unlike previous outcomes, the number of teen births in the twin family had no statistically significant effect, and the magnitude of the parameter was quite small, indicating that genetic or environmental influences shared by twin sisters are not responsible for any of the association between early motherhood and IP.

![Figure 3](attachment:figure3.png)

**Figure 3.** Effect of adolescent motherhood on Behavior Problems, when using different comparison groups. Regression coefficients for “Born to teen” (\(x_{1ijk}\)) appear in Table 4.
and IP. Figure 5 illustrates how the effect of being born to a teenage mother on IP remains approximately constant, even when comparing related children.

**Discussion**

Results from both descriptive means comparisons and HLMs are consistent with the hypothesis that adolescent motherhood causes increased risk for behavior, substance use, and internalizing problems in offspring. Specifically, offspring born to adolescent mothers demonstrate higher numbers of mental health problems than their later born siblings and higher numbers than their first cousins, comparisons that control for genetic and environmental variables shared by family members. Previous studies of the relation between adolescent motherhood and offspring adjustment that have failed to find significant associations when comparing first cousins (Geronimus et al., 1994; Turley, 2003) have primarily used young children, whereas the present study used adolescent and adult offspring aged 14–39 years. The adverse effects of adolescent motherhood for offspring mental health may be more evident as offspring themselves transition into adulthood, as has been previously suggested (Brooks-Gunn & Furstenberg, 1986). Overall, our results illustrate the usefulness of ostensibly “genetic” designs for elucidating putative environmental risks for psychopathology. Given our increased confidence that the relation between adolescent childbearing and offspring mental health

---

**Table 5**

*Hierarchical Linear Models of Adolescent Motherhood and Substance Use Problems*

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SE</td>
<td>M</td>
<td>SE</td>
<td>M</td>
</tr>
<tr>
<td>Random effects*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twin family - $\tau_{00k}$</td>
<td>6.14</td>
<td>1.73</td>
<td>5.80</td>
<td>1.55</td>
<td>5.71</td>
</tr>
<tr>
<td>Nuclear family - $\tau_{111}$</td>
<td>6.44</td>
<td>1.64</td>
<td>6.29</td>
<td>1.47</td>
<td>6.36</td>
</tr>
<tr>
<td>Children - $\alpha^2$</td>
<td>19.32</td>
<td>0.97</td>
<td>17.06</td>
<td>0.85</td>
<td>17.05</td>
</tr>
<tr>
<td>Fixed effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child gender*</td>
<td>2.83</td>
<td>0.29</td>
<td>2.83</td>
<td>0.29</td>
<td>2.81</td>
</tr>
<tr>
<td>Child age</td>
<td>1.65</td>
<td>0.25</td>
<td>1.66</td>
<td>0.25</td>
<td>1.63</td>
</tr>
<tr>
<td>Child age-squared</td>
<td>-0.03</td>
<td>&lt;.01</td>
<td>-0.03</td>
<td>&lt;.01</td>
<td>-0.03</td>
</tr>
<tr>
<td>Born to teen - $x_{1jk}$</td>
<td>1.93</td>
<td>0.59</td>
<td>1.69</td>
<td>0.56</td>
<td>1.51</td>
</tr>
<tr>
<td>First child as teen - $x_{0jk}$</td>
<td>0.46</td>
<td>0.43</td>
<td>1.52</td>
<td>0.99</td>
<td>-1.65</td>
</tr>
<tr>
<td>Teen births in twin family - $x_{00k}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent zygosity*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zygosity interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$-2LL$</td>
<td>8809.2</td>
<td>8667.2</td>
<td>8666.1</td>
<td>8661.4</td>
<td>8657.9</td>
</tr>
<tr>
<td>$[\Delta-2LL; \Delta df]$</td>
<td>[142; 3]</td>
<td>[1.1; 1]</td>
<td>[4.7; 1]</td>
<td>[3.5; 1]</td>
<td>0.31</td>
</tr>
</tbody>
</table>

*Note.* Boldface type: significant at $p < .05$.

*Variances of random effects estimated. *Female = 0; Male = 1. *monozygotic = 0; dizygotic = 1.

and IP. Figure 5 illustrates how the effect of being born to a teenage mother on IP remains approximately constant, even when comparing related children.

Figure 4. Effect of adolescent motherhood on Substance Use Problems, when using different comparison groups. Regression coefficients for “Born to teen” ($x_{1jk}$) appear in Table 5.
problems is truly environmental, however, the mechanisms of this environmental effect remain to be fully elaborated. How the effects of adolescent motherhood are perpetuated, even augmented, into the offspring’s adulthood is an intriguing problem.

Adolescent childbearing significantly predicted elevated mental health problems for children born when the mother was still an adolescent, but not for subsequent children born later in the mother’s life. This finding is consistent with Grogger (1997), who found that male children born to adolescent mothers had higher incarceration rates than their later born siblings. In contrast, Turley (2003) found that maternal age at first birth was more predictive of later born children’s cognitive test scores than maternal age at children’s own births. Moreover, Jaffee, Caspi, Moffitt, Belsky, and Silva (2001) found that maternal age at first birth significantly predicted early school leaving, unemployment, early parenthood, and violent criminal offending in young adult offspring, above and beyond maternal age at offspring’s own birth. To some extent, the discrepancy between the present results and these studies may be due to differences among the offspring outcomes in question, with cognitive or fertility outcomes attributable to selection effects to a larger extent than to mental health problems. However, violent criminal offending (Jaffee et al., 2001) is likely correlated with the behavior and substance use problems examined here. We may speculate that differences between the present study and Jaffee et al. (2001) reflect

### Table 6

**Hierarchical Linear Models of Adolescent Motherhood and Internalizing Problems**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Twin family - $\tau_{000}$</td>
<td>1.04</td>
<td>0.76</td>
<td>0.57</td>
<td>0.73</td>
<td>0.56</td>
</tr>
<tr>
<td>Nuclear family - $\tau_{111}$</td>
<td>2.94</td>
<td>0.91</td>
<td>3.05</td>
<td>0.89</td>
<td>3.05</td>
</tr>
<tr>
<td>Children - $\alpha^2$</td>
<td>12.15</td>
<td>0.59</td>
<td>11.86</td>
<td>0.57</td>
<td>11.86</td>
</tr>
</tbody>
</table>

#### Fixed effects

- Intercept: $3.78 \pm 0.14$
- Child gender: $-1.12 \pm 0.22$
- Child age: $0.17 \pm 0.19$
- Child age-squared: $<0.01 \pm 0.01$
- Born to teen - $x_{1ijk}$: $1.54 \pm 0.44$
- First child as teen - $x_{0jk}$: $0.18 \pm 0.30$
- Teen births in twin family - $x_{00k}$: $0.24 \pm 0.61$
- Parent zygosity: $0.19 \pm 0.33$
- Zygosity interaction: $-0.77 \pm 0.48$

#### Random effects

- Twin family: $-2LL = 7971.2 \pm [47.5; 3]$
- Nuclear family: $-2LL = 7923.7 \pm [3; 0.3]$
- Children: $-2LL = 7923.4 \pm [2.7; 1]$

Note. Boldface type: significant at $p < .05$.

*Variances of random effects estimated. Female = 0; Male = 1. Monozygotic = 0; dizygotic = 1.

---

**Figure 5.** Effect of adolescent motherhood on Internalizing Problems, when using different comparison groups. Regression coefficients for “Born to teen” ($x_{1ijk}$) appear in Table 6.
etiological differences between “adolescent-limited” and “life-course persistent” variants of antisocial behavior (Moffitt, 1993); however, a satisfactory explanation of this discrepancy is not really possible at this point and thus awaits future research.

Finally, being born into the type of extended family in which women gave birth as teenagers, even if not born to a teenager oneself, was predictive of higher levels of behavior and substance use problems. It appears that the children of adolescent mothers would be at elevated risk for behavior and substance use, even if their mothers delayed childbearing, because of background risk variables. These results make intuitive sense, given that adolescent mothers frequently have histories of conduct problems (Barone, Moffitt, Caspi, Dickson, & Silva, 1996; Woodward & Fergusson, 1999) and substance abuse (Ketterlinus, Lamb, & Nitz, 1994). Maternal antisocial behavior and substance use problems, in turn, are strongly predictive of similar problems in offspring (Frick & Loney, 2002; Jacob et al., 2003). In a longitudinal, intergenerational study, Emery, Waldron, Kitzmann, and Aaron (1999) found that delinquent behavior in childless adolescent girls predicted a lower age at first birth, nonmarital childbearing, and the subsequent behavior problems of their children. Measurement of parental psychopathology and personality is available for the Australian sample here, and future research should examine the extent to which transmission of externalizing problems accounts for the observed selection effects.

In addition, future research should attempt to quantify more precisely whether background variables confounding the association between adolescent motherhood and child behavior and substance use problems are genetic or environmental. Interaction parameter estimates from the hierarchical linear models were consistent with the conclusion that intergenerational associations with behavior problems and substance use problems reflect, in part, passive gene–environment correlation. However, these parameters were marginally nonsignificant. The children-of-twins design, like other behavioral genetic designs, demands extremely large samples to have adequate sample size. Given this demand, and the relative infrequency of adolescent childbearing in the present sample, the study may have lacked sufficient power to fully discriminate between environmental and genetic confounds. It will be important for future research to examine the underlying processes responsible for each outcome associated with teenage motherhood. Explicit quantification of HLM in terms of latent genetic and environmental components remains a difficult enterprise, although Rowe, Almeida, and Jacobson (1999) and McArdle and Prescott (2005) have described model fitting strategies useful in this regard.

Comparing Australia and the United States

As mentioned previously, the United States has one of the highest teenage pregnancy rates in the industrialized world (United Nations, 1991), primarily due to lower use of contraception and abortion, not different patterns of adolescent sexual activity or proportions of ethnic minorities. Approximately 5% of Australian adolescents experience a pregnancy each year (vs. 10% of U.S. adolescents each year), and over 50% of Australian adolescent pregnancies are aborted (vs. 35% of American adolescent pregnancies). Adolescent motherhood, therefore, is a considerably less common phenomenon in Australia. Perhaps because of the lower rate, the Australian government has not implemented any public policies designed to modify adolescent fertility or sexuality. The lack of institutionalized stigma surrounding adolescent childbearing, together with its cultural infrequency, may moderate the consequences for Australian children’s adjustment relative to American children. Despite quantitative differences in base rates, there is little evidence for qualitative differences between Australia and the United States in the biological, family, and sociocultural risk factors at play. Extensive research on the correlates of adolescent motherhood is not available for Australian populations (see Shaw et al., 2006); however, results from British and New Zealand populations are consistent with American research (Waldron, 2004).

It is important to note one essential difference between Australia and the United States: The non-Aboriginal population of Australia is almost exclusively Caucasian. This demographic difference is important in two respects. First, genetic variance is not a static characteristic of a trait; rather, it is a descriptive population statistic. As such, the magnitude of genetic variance in any phenotype may differ among various subpopulations (Shanahan & Hofer, 2005). This has been empirically demonstrated for a variety of phenotypes, including educational attainment across birth cohorts (Heath et al., 1985), physical health across income levels (Johnson & Krueger, 2005), childhood conduct problems across levels of family dysfunction (Button, Scourfield, Martin, Purcell, & McGuffin, 2005), alcohol use initiation across level of family religiousness (Koopmans, Slutske, van Baal, & Boomsma, 1999), and intelligence across socioeconomic status (Harden, Turkheimer, & Loehlin, 2007; Turkheimer, Haley, Waldron, D’Onofrio, & Gottesman, 2003) and across ethnic groups (Scarr, 1981; Willerman, 1979). Similarly, the relative magnitude of genetic variance in adolescent motherhood may differ among ethnic groups; this remains a possibility to be examined empirically.

Second, the existence of genetic influences on adolescent childbearing within a Caucasian subpopulation cannot be used to infer the origin of differences in adolescent childbearing rates between ethnicities. Geronimus (2004) has suggested that the higher rates of teenage pregnancy seen among African American and Latina women versus Caucasian women are driven entirely by cultural differences in perceived “optimal” fertility timing: Among women experiencing structural constraints to economic success, higher health and mortality risks, and a normative multigenerational family structure, early childbearing may be correctly perceived as advantageous. Future genetically informative research should use ethnically diverse American samples in order to examine potential cultural differences between the United States and Australia and between ethnic groups in the United States in the processes relating adolescent motherhood with child functioning.

Limitations in the Children-of-Twins Design: What About the Fathers?

The children-of-twins design is described as quasieperimental for good reason: Without random assignment of children to adolescent versus adult mother conditions, we cannot prove that adolescent childbearing does or does not cause child psychopathology. The children-of-twins design controls for environmental and genetic variables shared by twin sisters, but variables that vary systematically with adolescent motherhood within twin pairs remain as uncontrolled confounds. Although this list of potential

Comparing Australia and the United States

As mentioned previously, the United States has one of the highest teenage pregnancy rates in the industrialized world (United Nations, 1991), primarily due to lower use of contraception and abortion, not different patterns of adolescent sexual activity or proportions of ethnic minorities. Approximately 5% of Australian adolescents experience a pregnancy each year (vs. 10% of U.S. adolescents each year), and over 50% of Australian adolescent pregnancies are aborted (vs. 35% of American adolescent pregnancies). Adolescent motherhood, therefore, is a considerably less common phenomenon in Australia. Perhaps because of the lower rate, the Australian government has not implemented any public policies designed to modify adolescent fertility or sexuality. The lack of institutionalized stigma surrounding adolescent childbearing, together with its cultural infrequency, may moderate the consequences for Australian children’s adjustment relative to American children. Despite quantitative differences in base rates, there is little evidence for qualitative differences between Australia and the United States in the biological, family, and sociocultural risk factors at play. Extensive research on the correlates of adolescent motherhood is not available for Australian populations (see Shaw et al., 2006); however, results from British and New Zealand populations are consistent with American research (Waldron, 2004).

It is important to note one essential difference between Australia and the United States: The non-Aboriginal population of Australia is almost exclusively Caucasian. This demographic difference is important in two respects. First, genetic variance is not a static characteristic of a trait; rather, it is a descriptive population statistic. As such, the magnitude of genetic variance in any phenotype may differ among various subpopulations (Shanahan & Hofer, 2005). This has been empirically demonstrated for a variety of phenotypes, including educational attainment across birth cohorts (Heath et al., 1985), physical health across income levels (Johnson & Krueger, 2005), childhood conduct problems across levels of family dysfunction (Button, Scourfield, Martin, Purcell, & McGuffin, 2005), alcohol use initiation across level of family religiousness (Koopmans, Slutske, van Baal, & Boomsma, 1999), and intelligence across socioeconomic status (Harden, Turkheimer, & Loehlin, 2007; Turkheimer, Haley, Waldron, D’Onofrio, & Gottesman, 2003) and across ethnic groups (Scarr, 1981; Willerman, 1979). Similarly, the relative magnitude of genetic variance in adolescent motherhood may differ among ethnic groups; this remains a possibility to be examined empirically.

Second, the existence of genetic influences on adolescent childbearing within a Caucasian subpopulation cannot be used to infer the origin of differences in adolescent childbearing rates between ethnicities. Geronimus (2004) has suggested that the higher rates of teenage pregnancy seen among African American and Latina women versus Caucasian women are driven entirely by cultural differences in perceived “optimal” fertility timing: Among women experiencing structural constraints to economic success, higher health and mortality risks, and a normative multigenerational family structure, early childbearing may be correctly perceived as advantageous. Future genetically informative research should use ethnically diverse American samples in order to examine potential cultural differences between the United States and Australia and between ethnic groups in the United States in the processes relating adolescent motherhood with child functioning.

Limitations in the Children-of-Twins Design: What About the Fathers?

The children-of-twins design is described as quasiexperimental for good reason: Without random assignment of children to adolescent versus adult mother conditions, we cannot prove that adolescent childbearing does or does not cause child psychopathology. The children-of-twins design controls for environmental and genetic variables shared by twin sisters, but variables that vary systematically with adolescent motherhood within twin pairs remain as uncontrolled confounds. Although this list of potential
confounds is considerably shorter than traditional epidemiological studies, it still includes nonshared environmental influences on the twin mother and the genetic influences of the father (Eaves, Silberg, & Maes, 2005). Elevated rates of academic problems, school dropout, unemployment, delinquency, criminal behavior, and alcohol and substance use are seen in the fathers of children born to adolescent girls (Jaffee et al., 2001; Pears, Pierce, Kim, Capaldi, & Owen, 2005; Sigle-Rushton, 2005; Tan & Quinlivan, 2006). Men with genetic liabilities for, among other behaviors, delinquency and substance use are thus more likely to impregnate adolescent girls, and their children will likely inherit these same liabilities. Consequently, the higher numbers of behavior and substance use problems seen in the children of discordant twin pairs, a result superficially consistent with environmental causation, may at least partly be due to biological inheritance of paternal genetic liabilities. In other words, spousal genetic influence is confounded with the nonshared environmental variance in the twin parent, a notable limitation in using the children-of-twins design to examine the processes underlying associations between offspring functioning and any dyadic parental characteristic (Eaves et al., 2005). This limitation is mitigated, to some extent, by the inclusion of multiple children per twin parent. If siblings in the same nuclear family have the same father, then the test of whether the firstborn child of an adolescent mother fares worse than his or her siblings born later in the mother’s life controls for paternal genetic influence. Adolescent mothers, however, may be more likely than women who delay childbearing to have children by multiple partners, and thus differences among their children’s mental health problems may, at least in part, reflect inheritance of different parental genes. Unfortunately, we do not have sufficient information to ascertain the extent to which sibling offspring in our present study share paternity.

These limitations highlight several important points. First, it is critical for future research on adolescent childbearing to collect more information on fathers, even if they are nonresidential. This need is not limited to the present line of research but is common to research on families, especially high-risk families, in general. Second, controlling for the effects of genetic variation is a difficult task, even with complex behavioral genetic designs. Barring an improbable sample of female identical twins having children by male identical twins, disentangling the environmental effects of maternal age from paternal genetic effects will remain a considerable challenge. Despite these limitations, the present study provides more rigorous control of environmental and genetic background variables than any previous investigation, and thus represents a stride forward in researchers’ understanding of the processes by which adolescent motherhood is related to child functioning.

Conclusions

In the present American political landscape, few topics galvanize such acrimonious debates as adolescent sexuality. The trend toward earlier initiation of sexual intercourse has been characterized as, among other things, an inevitable consequence of secular decreases in pubertal timing (e.g., Parent et al., 2003), the liberation of youth from antiquated sexual mores (e.g., Reiss, 1990), or an indication of catastrophic moral decline (e.g., Popenoe, 1998). Unlike Western European nations, where delegating to public schools the task of educating adolescents about reproductive biology, contraceptive use, and sexually transmitted diseases is little contended, factions in the United States continuously debate not only what sexual education curricula should entail but also whether such education should be the proprietary domain of the family rather than the schools. Perhaps not surprisingly, these debates have extended to adolescent childbearing: Many researchers and policymakers portray the lives of adolescent mothers and their children as extremely bleak (e.g., Hayes, 1987), whereas others consider the idealization of delayed fertility as indicative of ethnic bias and emphasize potential advantages to early childbearing (e.g., Geronimus, 2004). The present results, however, are consistent with a more nuanced view of adolescent motherhood. A narrow focus on delaying fertility, without ameliorating the background risks experienced by adolescent mothers and their children, may not be fully effective in reducing offspring behavior and substance use problems. Far from being innocuous, however, adolescent motherhood was associated with elevated risks for behavior problems, substance use problems, and internalizing problems in offspring, suggesting that delaying childbearing until adulthood may have long-term positive effects on the psychological adjustment of offspring. Given the large number of children born to adolescent mothers in the United States (422,197 births to teenagers in 2004 alone; Child Trends, 2006), policies or interventions effective in reducing the adolescent pregnancy rate may have widespread benefits for public mental health.

References

Button, T. M. M., Scourfield, J., Martin, N., Purcell, S., & McGuffin, P.
Rutter, M., Picketts, A., Murray, R., & Eaves, L. (2001). Testing hypotheses...

Received May 1, 2006
Revision received March 12, 2007
Accepted April 19, 2007