Published in final edited form as:

Behav Genet. 2012 January ; 42(1): 30–39. doi:10.1007/s10519-011-9481-2.

Observed Externalizing Behavior: A Developmental Comparison of Genetic and Environmental Influences Across Three Samples

Kristine Marceau,

The Pennsylvania State University, University Park, PA, USA

Mikhila N. Humbad,

Michigan State University, East Lansing, MI, USA

S. Alexandra Burt,

Michigan State University, East Lansing, MI, USA

Kelly L. Klump,

Michigan State University, East Lansing, MI, USA

Leslie D. Leve, and

Oregon Social Learning Center, Eugene, OR, USA

Jenae M. Neiderhiser

The Pennsylvania State University, University Park, PA, USA

Kristine Marceau: kpm170@psu.edu

Abstract

Estimates of genetic and environmental influences on externalizing behavior are markedly inconsistent. In an attempt to refine and extend our knowledge of externalizing behavior, the current study examined the etiology of externalizing behavior using observational data in middle childhood and adolescence from three twin and sibling samples. Observational ratings offer a unique perspective on externalizing behavior rarely examined within behavioral genetic designs. Shared environmental influences were significant and moderate to large in magnitude across all three samples (i.e., 44, 77, and 38%), while genetic influences (31%) were significant only for the adolescent sample. All three samples showed greater shared environmental influences and less genetic influence than is typically found when examining self-, parent-, and teacher-reports of externalizing behavior. These findings are consistent with other reports that have found evidence for shared environmental influences on measures of child externalizing behavior—in direct contrast to a commonly held perception that shared environmental factors do not have significant influences on behavior beyond early childhood.

Keywords

Externalizing; Genetic; Observational data

There is widespread interest in understanding the causes of externalizing behavior problems. Genetically-informed studies consistently indicate that the majority of variance can be explained by genetic influences with little contribution of shared environmental influences, although some reports indicate significant, sizable shared environmental influences (Burt

2009a). There are several explanations for these differences across studies, including definition specificity, age, and error. The measurement of externalizing behaviors offers yet another compelling possibility, as it is now widely acknowledged that heritability estimates vary by informant (Burt 2009a). Though researchers investigating externalizing behavior using non-genetically informed designs make good use of a variety of assessment methods, including parent-, teacher-, self-, and observer-reports, very few behavioral genetic studies have assessed behavior through observer ratings. This is potentially problematic as informant-reports can be affected by the dispositional characteristics of the informant and possible sources of bias (e.g., maternal rater contrast effects, unreliability in child self-report; see Burt 2009B). Observer-rated information offers a unique window into behavior because an individual's actual behavior is recorded in real time. As such, observational data augment self- and other-reported data, creating a more complete picture of the etiology of externalizing behavior. Accordingly, it is important to incorporate observer reports into genetically informed studies. The focus of the present study is to examine genetic and environmental influences on observational ratings of externalizing behavior within three different genetically-informed samples of twin children during middle childhood to adolescence.

Importantly, genetically informative samples generally represent normatively developing youth, not at risk samples, and the current study is no exception. Thus, the range of externalizing behaviors in these studies is normative, and only few children approach subclinical and clinical levels of externalizing problems. That said, examining the genetic and environmental influences on externalizing behaviors is important for understanding the reasons typical children display these behaviors, and the results from such studies should help to inform researchers studying at risk populations (for example, by contrasting the etiology of elevated problems with normative behaviors).

Variability in heritability estimates

Understanding the sources of the differences found in estimates of genetic and environmental influences on externalizing behavior across studies can improve the understanding of the etiology of externalizing behaviors. Variation in these estimates may occur because genetic and environmental influences on problems differ because of sample characteristics (e.g., age, SES, sex), or because the measurement is imprecise. The former represents meaningful variability, whereas the latter represents a failure in methodology. In a recent meta-analysis, Burt (2009a) aimed to quantitatively study how age, sex, measuring instrument, and sampling error may contribute to variability in the estimates of genetic influence on aggressive and non-aggressive antisocial behavior. She concluded that although heritability and shared environment both explain individual differences in aggressive and non-aggressive antisocial behavior, operationalization, age, and assessment method moderated the relative influences of genes and environment, consistent with findings from another meta-analysis, Rhee and Waldman (2002). Increasing the understanding of how such factors influence estimates of genetic and environmental influences on externalizing problems can provide insight into the assessment methodologies, facilitate efforts to increase the accuracy of estimates, and inform measurement of future studies.

Definitional specificity

Estimates of genetic and environmental influences differ according to the set of externalizing behaviors studied (i.e. aggression versus delinquency; and conduct disorder versus hyperactivity) (e.g. Van der Valk et al. 1998; Dick et al. 2005). When discussing normative externalizing behaviors, distinctions are often made between aggressive and non-aggressive behaviors. Aggressive behaviors typically refer to physical, often violent behaviors, whereas non-aggressive behaviors include delinquency, and rule breaking. When

non-aggressive and aggressive antisocial behavior were considered separately in a recent meta-analysis, genetic influences accounted for approximately half of the variance in non-aggressive antisocial behavior (48%) with the remaining being split between shared and nonshared environmental influences (Burt 2009a). In contrast, for aggression, additive genetic influences (65%) and nonshared environmental influences (30%) accounted for most of the variance, leaving little explained by shared environmental influences (5%). Thus, shared environmental influences explained more of the variance in non-aggressive externalizing behavior than aggressive externalizing behaviors (see also Rutter et al. 1990b; Van den Oord et al. 1994).

Age

Another source of heterogeneity in estimates of genetic and environmental influences on externalizing behaviors is the age of the individual. Overall, most studies have shown that genetic and nonshared environmental influences increase, whereas shared environmental influences decrease from adolescence to adulthood (i.e., as individuals mature and widen social circles beyond the home; Miles and Carey 1997). Burt's (2009a) meta-analysis grouped youth into three age groups: 1–5, 6–10, and 11–18. Results examining age showed that genetic and environmental influences on aggressive versus non-aggressive behaviors did not differ in early and middle childhood, but that by adolescence aggressive behaviors demonstrated greater genetic influences while nonaggressive behaviors demonstrated greater shared environmental influences. Moreover, genetic influences on aggression increased with age while shared environmental influences decreased, but for nonaggressive behaviors, genetic influences decreased with age while shared environmental influences remained stable (Burt 2009a). Given these results, the present study focuses on middle childhood and adolescence in an attempt to clarify the relative importance of genetic, and especially shared environmental influences, on non-aggressive externalizing behaviors during this developmental period.

Measurement

Heritability estimates have also been found to fluctuate according to measurement and informant. For example, nonshared environmental influences were greater for teacher ratings of externalizing behaviors than for parent-reported behaviors, whereas genetic influences were greater in parent-reported behaviors than teacher-reported behaviors (Towers et al. 2000). Rhee and Waldman (2002) also showed that nonshared environmental influences were greater for self-reported externalizing behaviors than other-reported externalizing behavior, whereas genetic influences were greater for externalizing behaviors reported by others than for self-reported behavior. They, however, were unable to compare estimates of genetic and environmental influences from observer ratings of externalizing behavior to self- and other-reported behavior because too few studies used observer reports (Rhee and Waldman 2002).

There are relatively few studies examining observer ratings of youth externalizing behavior (Burt et al. 2011; Leve et al. 1998; Plomin et al. 1981; Rende et al. 1992; O'Connor et al. 1995). In middle childhood, shared environmental influences contributed to a large proportion of the variance in externalizing behavior (e.g. Leve et al., 1998; Plomin et al. 1981; Rende et al. 1992). In a sample of 10–18 year olds, shared environmental influences explained 31% of the variance in 'acting out' behaviors (Burt et al. 2011), consistent with previous reports in this age range (e.g. O'Connor et al. 1995).

One strength of observer reports is that behaviors can be defined consistently and reliably by the researcher instead of relying on individual interpretations of a definition likely to be specific to the individual parent, child, or teacher reporter (Gardner 2000). Observer ratings

may reduce rater bias, circumventing parental expectations or contrast effects, and like teacher-reports, allows the examination of youth behavior on a more specific level, narrowing definitional specificity. Further, children are all rated in comparison with other children of the same age by a single set of raters for the whole sample who are not genetically-related to the children. Using observer reports can therefore add to our knowledge of the etiology of externalizing behaviors by examining the specific behaviors of interest in a controlled setting, giving us information from a different perspective than parent-, teacher-, and self-reports.

The current study

Using three samples assessed during middle childhood and adolescence, we examined genetic and environmental influences on non-aggressive externalizing behaviors. By examining only observer reports, we hope to help clarify genetic and environmental influences on specific externalizing behaviors. Based on the above review, we specifically hypothesized that shared environmental influences will be a significant source of variance in observer-rated externalizing behavior regardless of age, but will be especially prominent in the middle-childhood samples, and that genetic influences will be greater in the adolescent sample relative to the middle childhood samples.

Methods

Michigan State University (MSU) Twin Registry

Participants—The MSU sample consisted of 100 twin families assessed as part of the ongoing Twin Behavior and Emotional Development-Children (TBED-C), one study in the Michigan State University Twin Registry (MSUTR). Twin families were recruited via State of Michigan birth records in collaboration with the Michigan Department of Community Health. Zygosity was established using physical similarity questionnaires that show 95% accuracy or better (Peeters et al. 1998), via telephone prior to the family's assessment. During the assessment, a research assistant independently evaluated twins on physical similarity indices. Unclear or discrepant zygositys were resolved through DNA markers (see Klump and Burt 2006).

Twin pairs were 6–10 years old ($M = 8.3$; $SD = 1.3$); 41 monozygotic (MZ) and 59 same-sex dizygotic (DZ) twin pairs. Participating families endorsed ethnic group memberships, parental education, and poverty at rates comparable to those of other area inhabitants (Culbert et al. 2008). See Table 1 and Klump and Burt (2006) for additional demographic information.

Observed externalizing behavior—Each mother–child dyad completed a mildly-to-moderately frustrating 8-minute task in an office space restructured to resemble a living room. Dyads used an Etch-a-Sketch to draw specific pictures, but each member of the dyad could use only one dial, thereby requiring cooperation within the dyad. Interaction data were coded using the twin parent–child interaction system (PARCHISY) (Deater-Deckard et al. 1997). Both the task and the coding system are reliable and valid tools with school-age children (Deater-Deckard and Petrill 1999). Separate staff coded each sibling's interaction to minimize rater bias effects. Research assistants were blind to zygosity status and informant-reports of child externalizing. The order of participation was counterbalanced for birth order of the siblings.

The current study examined an averaged composite of three scales, each tapping different aspects of externalizing behavior and scored on a 7-point, Likert-type scale: child non-compliance (e.g., refusal to follow parental requests/commands), child negative affect (e.g.,

frowning, cold/harsh voice), and child on-task behavior (e.g., initiative, persistence with regard to assigned task; reverse-scored). 15% of videotapes were reviewed by an independent rater to determine reliability. Inter-rater reliability was acceptable for each of the three coded behaviors (ICCs > 0.85 in all cases).

Oregon Twin Project (OTP) sample

Participants—The OTP sample consisted of 150 twin families identified between 1993 and 1994 via birth announcements, twin organizations, and the public school system in Oregon. Zygosity was determined by three raters using the primary caregiver's report on the Zygosity Questionnaire (Goldsmith 1991), which taps similarity of physical, developmental, and medical conditions and photographs. Zygosity questionnaires have been shown to be about 95% accurate (Goldsmith 1991). One twin pair was excluded from analyses because they could not be reliably classified.

Twins were 7–13 years old ($M = 10$ years, 2 months; $SD = 22$ months): 77 MZ twin pairs and 72 DZ twin pairs (31 were male/female pairs). See Table 1 and Leve (2001) for additional demographic information.

Observed externalizing behavior—Twin-friend dyads participated in a 10-min videotaped interactive task planning a fun activity (5 min) and talking about the next school year (5 min) during a laboratory visit. Each friend was the same sex, age and/or grade as the twin, genetically unrelated to the twin, living in a separate household, and a different friend than their co-twin's friend. Following the task, a research assistant provided global ratings of child behavior using a 5-point, Likert-type scale. All research assistants were blind to participant zygosity. The order of participation was counterbalanced for birth order. Separate staff coded each sibling's interactions. An observed externalizing score was formed from 12 global rating items that indicated the degree of each twin's antisocial, argumentative, and uncooperative behavior during the interactions, $\alpha > 0.91$ (see Leve, 2001). Example items include 'Did the target child initiate arguments?' 1 = never, 5 = often', and 'rate the target child on the following bipolar adjectives (e.g. rude—polite, unpleasant—pleasant, uncooperative—cooperative) using the following scale, 1 = very, 2 = somewhat, 3 = neutral, 4 = somewhat, 5 = very'. Interrater reliability was calculated by the percent agreement on 15% of the cases and indicated good reliability (percent agreement ranged from 92% to 97% across raters).

Nonshared Environment in Adolescent Development (NEAD) study sample

Participants—The NEAD sample consisted of 720 families. A subsample of families was recruited through random digit dialing of 10,000 telephone numbers throughout the United States; however, most of the families were recruited through a national market survey of 675,000 families (Reiss et al. 2000). Families included MZ twins (92), DZ twins (94), and full siblings (FI; 90) in non-divorced families, and full siblings (FS; 173), half siblings (HS; 105), and genetically unrelated siblings (US; 124) in stepfamilies. To establish zygosity, twins were rated for physical similarity using a questionnaire designed for adolescents (Nichols and Bilbro 1966). Approximately 6% of the twin sample could not be classified with certainty and were excluded from further analyses.

The sample consisted of same-sex siblings aged 10–18 years old ($M = 13.6$; $SD = 3$ years). Siblings were all within 4 years of age of each other and lived in the same household for at least 5 years. Finally, both siblings were required to live in the same household at least half of the time. See Table 1 and Reiss et al. (2000) for sample demographic information.

Observed externalizing behavior—Families were visited in their homes, where children and parents completed questionnaires. Family members were videotaped in three 10-min dyadic (mother–child, father–child, and sibling–sibling) problem-solving discussions (topics were identified by questionnaires completed prior to the visit). For each interaction, observers provided global ratings of the child’s antisocial behavior (Hetherington and Clingempeel 1992). The antisocial behavior scale measured the degree to which the child disrupted the interaction or was disrespectful towards authority or peers on a scale of 1 to 5 (1 = no antisocial behavior present, 2 = minimal evidence of antisocial behavior, behavior abated quickly, 3 = occasional display of low intensity antisocial behavior less quickly abated, 4 = moderately intense antisocial behavior, 5 = frequently demonstrated antisocial behavior). Coders were instructed to code for rude, inconsiderate, noncompliant, uncooperative, irritable, hostile, coercive or aggressive behavior. A measure of observed externalizing behavior was created by averaging the global observer ratings of all three interactions. Research assistants were blind to zygosity, and the order of participation was counterbalanced for birth order. Separate staff coded each sibling’s interaction. The intra-class correlations for observed antisocial behavior across the three situations were above .79. The reliability of coders for each sibling’s antisocial behavior in each situation was also acceptable, kappas > 0.65, mean ICC > 0.86. See Henderson (1999) for a detailed description of the observational coding scheme.

Statistical approach

Because the individual samples consisted of males and females spanning a wide age range, we regressed age, sex, and age differences (for nontwin sibling pairs) out of externalizing scores within each sample. Standardized residuals were used for all subsequent analyses (McGue and Bouchard 1984). Thus, scores for each of the three samples were standardized to have a mean of zero and a standard deviation of one. No significant mean-level differences in observed externalizing behaviors were found between sibling types in any of the three samples.

Data analyses proceeded in two steps. First, means and standard deviations were examined across sibling type (monozygotic twins (MZ), dizygotic twins (DZ), full siblings (FI, FS), half siblings (HS), and genetically unrelated step siblings (US) and sibling intraclass correlations (ICCs) were computed using double-entered data separately for each sibling type for each sample (correlating sibling 1 with sibling 2). Genetic influences are suggested if ICCs decrease according to decreasing genetic similarity ($MZ > DZ = FI = FS > HS > US$). Shared environmental influences are suggested if ICCs are similar for genetically non-identical siblings ($DZ = FI = FS > 0$, $US > 0$). Finally, MZ ICCs < 0 indicate non-shared environmental influences.

Biometrical models of the standardized raw data were next examined using Mx (Neale 1999) to systematically estimate genetic and environmental influences. We tested two models, an unconstrained model allowing the genetic and environmental estimates to vary by sample, and a constrained model, wherein paths were set to be equal across samples. If the constrained model could be accepted, then we have no statistical evidence that the estimates are unequal across samples. Model differences were explored using a nested model approach that compared the constrained to the unconstrained model. Differences in chi-square estimates between models tested whether the constrained model resulted in a significant deterioration of fit. The constrained model was accepted as the best fitting and most parsimonious if no significant deterioration of fit was found.

In these models, the correlation between the genetic factors (A) was set to equal the siblings’ degree of genetic similarity. The correlation between the shared environmental factors (C) was set to be 1.0, as shared environmental influences are defined as non-genetic influences

that make siblings who are reared in the same family similar to one another. Finally, the correlation between the nonshared environmental factors (E) was fixed to 0 for all relative pairs; by definition, such influences make siblings different from one another. E also includes measurement error. The overall fit of the model was tested by -2 times the log-likelihood of data ($-2\ln L$) and the Akaike's Information Criterion (AIC; Akaike 1987). A nonsignificant $-2\ln L$ and χ^2/df ratios > 2 indicate a good model fit; lower AIC values also indicate better model fit.

Model assumptions—Univariate biometric models are built on the assumption that shared and nonshared environmental effects are the same across sibling types and that there is no assortative mating (i.e., nonrandom mating that could result in suppressed estimates of genetic influences because DZs and full siblings would share more than 50% of their genes). No systematic differences were found for the validity of equal twin and sibling environments in NEAD (Reiss et al. 2000), suggesting that the equal environments assumption is tenable for the current study. Sizable correlations between spouses for the same characteristic are indicative of assortative mating. These effects cannot be fully tested in this report because individual antisocial behavior data were not available on both parents in any of the three studies. Assortative mating has generally been found to be modest for psychological traits (e.g. Plomin et al. 1990), with the exception of antisocial behavior (e.g. Du Fort et al. 2002). If assortative mating is operating for genes related to externalizing behavior in youth, shared environmental influences on the phenotype may be inflated. However, in the NEAD sample, the net effect of assortative mating is reduced by the inclusion of genetically unrelated siblings (Pike et al. 1996).

Results

Intraclass twin/sibling correlations for observed externalizing

The intraclass twin/sibling correlations are presented separately by sample in Table 2. As evidenced in the table, DZ twins had equivalent correlations to MZ twins for the MSU and OTP samples, suggesting substantial shared environmental influences and modest nonshared environmental influences on observed externalizing behavior in these samples. In the NEAD sample, however, the slight cascade in correlations across sibling types indicated some genetic, shared, and nonshared environmental influence.

Model-fitting results

The constrained model (constraining all genetic, shared, and nonshared environmental influences across all three samples) showed a significant decrement in model fit, $X^2\text{change}(4) = 30.7$, $p < 0.05$ (unconstrained model: $-2\ln L(1840) = 4872.3$, AIC = 1193.2; constrained model: $-2\ln L(1846) = 4903.9$, AIC = 1211.9). Thus, the unconstrained model provided the better fit to the data, as evidenced by a significant increase in chi-square when constraining parameter estimates to be equal and the higher AIC value, suggesting that parameter estimates do vary significantly across samples. Parameter estimates for the full and best fitting models for each sample are presented in Table 3.

The MSU sample showed no evidence of genetic influence, and substantial shared and nonshared environmental influence (although only the latter was significantly greater than zero, which likely reflects the small sample size). The OTP sample also showed no genetic influence, a very large shared environmental influence, and modest nonshared environmental influences. The NEAD sample showed significant genetic influences, shared environmental influences, and nonshared environmental influences.

Because there was no evidence of genetic influences in the MSU and OTP samples, the model was re-run dropping genetic influences on observed externalizing behaviors in these two samples (i.e., $-2\text{Lnl}(1842) = 4873.2$, $\text{AIC} = 1189.2$). Unsurprisingly, the C and E estimates did not change when dropping the A path. However, the non-shared environment accounted for slightly more of the variance in the MSU sample than did the shared environment (E: 56%, 95% CI: 38–74%; C: 44%, CI: –58 to 58%). For the OTP sample, shared environment explained the vast majority of the variance in observed externalizing behavior (77%, CI: 61–83%), with the remaining variance attributable to E (23%, CI: 17–30%). There were approximately equal genetic influences (31%, 95% CI: 14–46%), shared environmental influence (38%, 95% CI: 28–48%), and nonshared environmental influences (30%, 95% CI: 22–41%) in the NEAD sample. Although the three samples differed in their heritability estimates, it is clear from the results that shared environmental influences are important for all three samples.

Discussion

Although there is a substantial body of research on the etiology of externalizing behaviors, our understanding of etiological processes is far from complete. Prior studies using a single sample and/or a single reporter (parent, self, and/or teacher) of externalizing behavior have resulted in an inconsistent pattern of findings regarding the relative influences of genetics, shared environment, and nonshared environment. In the present study, we aimed to advance the understanding of the etiology of externalizing behavior as measured via observational assessment. For a complete understanding of genetic and environmental influences on externalizing behaviors, these behaviors must be examined both globally and in more specific situations. We utilized a multisample, observational approach in an effort to reduce sample-dependence and to maximize our ability to estimate genetic and environmental influences on observational ratings of externalizing behavior, the measurement strategy least represented in genetically informed literature.

There were significant differences in estimates of genetic and environmental influences across samples. Overall, the results indicated that shared environmental influences are important for observed youth externalizing behavior, with the remaining variance accounted for by nonshared environmental influences, and, in one study, genetic influences. Each study was designed to tap into a slightly different specific context, thus the interpretation of the pattern of results is more generalizable than if only one specific context was examined.

Quantitative behavioral genetic designs allow us to disentangle the roles of genetic and environmental factors for specific behaviors. Attempts to identify specific nonshared environmental influences have been mostly unsuccessful (Neiderhiser et al. 2007), possibly because nonshared environmental influences are time-specific and/or idiosyncratic (Burt 2009b; Turkheimer and Waldron 2000; Rutter et al. 1999). Shared environmental influences, however, appear to be both identifiable and persistent over time (at least through childhood and adolescence; Burt 2009b). These findings suggest that observer reports may be especially insightful when investigating specific shared environmental effects. Observers may be particularly sensitive to shared environmental influences on behavior. Thus, using observer reports during the developmental period when shared environmental influences potentially show the greatest influence on behavior may help us to detect shared environmental influences, despite the power issues often experienced in twin studies that limit the ability to detect shared environmental influences. Observer reports may also be unbiased by genetic influences on the perceptions of externalizing behavior carried by parents or youth.

There is some evidence that observer reported behavior may better predict later behavior than parent and teacher reported behavior. For example, observer reported externalizing behavior in childhood, but not parent reported behavior predicted later arrest rates (Patterson and Forgatch 1995). Observed negative behaviors but not reported behavior in childhood predicted later teacher reported externalizing behavior at age 5 (Fagot and Leve 1998). Applied to genetically informative studies, using observer reports may increase our ability to identify specific environmental influences underlying these associations between childhood externalizing behaviors and later problems.

Measurement issues when studying externalizing behavior

Integrating observed and reported externalizing behaviors is also important for consistency of findings across fields. Despite differences between estimates across samples, when examined relative to the existing literature, this study provides evidence that the environment plays a greater role in observed behavior than in self- or other-reported behavior (as suggested in existing meta-analyses; see Burt 2009a; Rhee and Waldman 2002). For example, in the OTP sample, there were nonsignificant estimates of shared environmental influences on child-reported and teacher-reported externalizing behavior, while the estimates of genetic (39%), shared environmental (46%) and nonshared environmental (15%) influences on parent-reported externalizing behavior were all significant (Bullock et al. 2006). In contrast, observer reports presented in the current study showed moderate to substantial shared and nonshared environmental influences and no genetic influences.

In the NEAD sample, there were nonsignificant estimates of shared environmental influences on mother- (6%) and teacher- (0%) reported externalizing behavior, but significant estimates of shared environmental influences on father-reported (23%) externalizing behavior, and in estimates using observer reports in the present study (Deater-Deckard et al. 1997; Towers et al. 2000). Estimates of genetic influences on mother- (83%) and father- (59%) reported externalizing were much higher than estimates of genetic influences on teacher- (31%) reported externalizing and estimates obtained using observed externalizing behavior in the present study (Deater-Deckard et al. 1997; Towers et al. 2000). Specific data on the genetic and environmental influences on externalizing behavior rated by parents or youth are not yet available for the MSU sample.

Our estimates of genetic effects accounted for far less than half of the variance in the current study, results that stand in striking contrast to those using other informant-reports of these behaviors (e.g., Van der Valk et al. 1998; Leve et al. 1998; Van den Oord et al. 1994). Moreover, our estimates of shared environmental influences were on the high end of, and in one case exceeded, the range of shared environmental influences on externalizing behaviors during middle childhood noted in prior reviews (0–62%) (Deater-Deckard and Plomin 1999). This enhancement of the estimation of shared environmental influences and the corresponding reduction in genetic influences when using observational measures has been noted using different measures and measurement occasions for two of the samples used in this report (Leve et al. 1998; O'Connor et al. 1995). This is consistent with the few available studies that have yielded evidence of stronger shared environmental influences and lower estimates of genetic influences than questionnaire data in child maladaptive behavior and adolescent behavior towards parents (e.g., Burt et al. 2011; Leve et al. 1998; Miles and Carey 1997; O'Connor et al. 1995; Plomin et al. 1981; Rende et al. 1992).

Differences in estimates of genetic and environmental influences by method of assessment might arise from a number of factors, including parental expectations (e.g., that identical twins will be more similar to each other than fraternal twins, inflating estimates of genetic influences), rater contrast effects (e.g., parents rating two children may contrast the siblings

in their ratings, thus magnifying differences, inflating estimates of nonshared environmental influences), and differences in the underlying construct tapped by the different methods. Parent- and youth-report offer a window into more broad-based, trait-like levels of externalizing behavior that a parent may remember over the course of the last 6-months, or the entirety of a child's life. While this information on externalizing behaviors broadly is useful, parent- and youth-reported scores may be colored by the rater's own personality and/or psychological issues (e.g., a child's externalizing behavior may be particularly salient and overestimated, or a depressed mother may not have an accurate view of her child's behavior).

Together, findings suggest the importance of a multi-method approach for understanding the etiological sources of externalizing behavior. Studies investigating the predictive validity of composite scores using multiple reporters vs. single reporters typically show that composites across raters perform better than single raters, however they also generally exclude observer ratings (e.g. van Dulmen and Egeland 2011). Because observer ratings provide an additional, different perspective on child behavior, studies would be strengthened by also including observer reported behavior in composites. Interestingly, when composites across multiple raters, including observer ratings, have been used, heritability estimates tend to be *higher* than for any one rater (e.g. Reiss et al. 2000; Arseneault et al. 2003). Present findings corroborate limited existing evidence suggesting that global measurement shows greater genetic influences, whereas measurement in more specific contexts may show greater environmental influences (i.e. teacher reports, e.g. Towers et al. 2000, and observer reports). Parent and youth reported externalizing behaviors may tap into more global contexts, whereas teacher and observer reported behaviors are used to tap into more specific contexts. Thus, observer and parent-/youth-/teacher-reported data can each help further the understanding of different aspects of externalizing behavior.

It is important to note the limitations of observer ratings. The specific context may be artificial, and covers only a short time (e.g. 10 min). Therefore, the scores given may change dramatically because of short term fluctuations in behavior. In addition, the observed behavior may be specific to the interaction (e.g. parent-child, sibling, or friend). Finally, though observer reports are thought to reduce rater bias because of extensive and continued reliability training (Gardner 2000), we cannot know whether there are in fact biases present for raters, and whether these potential biases differ systematically across studies. Thus, we are not advocating for the sole use of observer reports, rather, we argue that a more complete understanding of etiology can be gleaned by taking advantage of the positive aspects of all assessment methods available. It should also be noted that observational ratings of specific behavior in older youth and adolescents may be more difficult to interpret because behaviors are observed during verbal, dyadic interactions and often tap into behavior in the context of a relationship.

Variability in the heredity of externalizing behaviors

By limiting our study of externalizing behavior to middle childhood and adolescence, and only examining observed behavior, the present study indicated that the types of externalizing behavior that can be observed during dyadic interactions tend to be more influenced by shared environment than are general recollections of externalizing behaviors as summarized in the existing literature. However, even limiting the present study to observationally-assessed externalizing behavior, we still found cohort differences in estimates of genetic and environmental influences.

Sample differences—The genetic contribution to externalizing behavior was greater for the sample with the oldest youth (NEAD), and shared environmental influences were large

in the OTP sample, while moderate in the other sample. There is some evidence that genetic and environmental influences on antisocial behavior will vary as a function of sample age and timing of assessment in development (Rhee and Waldman 2002; Miles and Carey 1997). It is possible that the age of the sample is driving the differences in the heritability estimates between samples, though we were unable to test whether differences were attributable to age because there was not sufficient power in the MSU and OTP samples.

It is also possible that the differences found are cohort effects not attributable to age, but rather to differences in the observational assessment methodologies used across the three studies. The NEAD sample used nontwin siblings while the other samples contained only twins. However, estimates of genetic and environmental influences did not differ significantly when examining only twins from the NEAD sample versus the entire NEAD sample (data not presented, available from the author upon request). Thus, since estimates still differed when comparing twins from NEAD with twins from the other samples, it is unlikely that sample differences were a result of sample type. However, the differences in estimates across samples could be attributable to geographical differences or systematic differences in raters across studies.

Further, although the measures analyzed in the three samples were conceptually similar and followed a similar methodological approach, the exact measures differed. Differences in scoring protocols could explain differences between samples. For example, MSU used micro-coded episodes, whereas OTP and NEAD used global coding. Though all three samples used dyadic interactions, one interaction was with the mother, one with a friend, and one was a composite of interactions with the mother, father, and sibling. While we consider this a strength for this study because it increases the generalizability of the findings, it also makes comparisons across the studies more difficult and may have contributed to differences in estimates across samples. Youth may behave differently with each of these partners in the dyad. Such variance in situation and coding might have contributed to the cohort differences observed.

Limitations and future directions

Although the multisample approach and including non-twin siblings strengthens the current study by providing increased power and increased generalizability of the results, this approach has some limitations. In addition to sample differences already discussed, it is also possible that we have inflated estimates of shared environmental influences because of shared method variance. That is, both twins/siblings are put into the same situation, and in the case of two of the samples, discussing similar topics with the same parent.

Further, we were unable to account for genotype–environment correlation processes shaping the discussion, and the child’s behavior in the discussion. Parents may provide both a common stimulus for the twins and the same genotype–environment correlation processes may influence these discussions. This may be especially salient if coming from both parents (i.e. if there was assortative mating for antisocial behavior, which we were unable to test). This limitation also applies to twin-friend interactions. Twins may share friends in middle childhood, and because youth pick their own friends frequently on characteristics for which the children are similar (Rowe et al. 1994), there may be similarities in the friend that each twin participated with that also inflate shared environmental influences in the third sample.

Despite these limitations, the present study is an important step in understanding the etiology of observed externalizing behaviors. In the face of methodological and age differences among three samples, observer ratings yielded larger shared environmental influences and smaller genetic influences than previously-reported parent- and youth-reported data have shown. Observer ratings are a valuable resource for examining behavior

that is largely ignored in behavioral genetic literature, and future studies would benefit from integrating observer reported data with parent-, youth- and teacher-reported data in order to take full advantage of all available methodologies and to further the understanding of the etiology of externalizing behavior. Likewise, future studies searching for specific shared environmental influences might increase their likelihood of success by taking advantage of the increased influence of the shared environment in observed data.

Acknowledgments

The MSU project was funded by the Intramural Grants Program, Michigan State University, #04-IRGP-232 and R01 MH081813. The Oregon Twin Project was funded by P30 MH 46690 and R03 MH 57053. Other research funds were provided by the University of Oregon, Department of Psychology. Data collection and other work on the Nonshared Environment in Adolescent Development project was supported by R01 MH43373, R01 MH48825, R01 MH059014, R01 MH065563 and the William T. Grant Foundation.

References

- Akaike H. Factor analysis and AIC. *Psychometrika*. 1987; 52:317–332.
- Bullock BM, Deater-Deckard K, Leve LD. Deviant peer affiliation and problem behavior: a test of genetic and environmental influences. *J Abnormal Child Psychol*. 2006; 34:29–41.
- Burt SA. Are there meaningful etiological differences within antisocial behavior? Results of a meta-analysis. *Clin Psychol Rev*. 2009a; 29:163–178. [PubMed: 19193479]
- Burt SA. Rethinking environmental contributions to child and adolescent psychopathology: a meta-analysis of shared environmental influences. *Psychol Bull*. 2009b; 135:608–637. [PubMed: 19586164]
- Burt SA, Klahr AM, Rueter MA, McGue M, Iacono WG. Confirming the etiology of adolescent acting-out behaviors: an examination of observer-ratings in a sample of adoptive and biological siblings. *J Child Psychol Psychiatr*. 2011; 52:519–526. [PubMed: 20955188]
- Culbert KM, Breedlove SM, Burt SA, Klump KL. Prenatal hormone exposure and risk for eating disorders: a comparison of opposite-sex and same-sex twins. *Arch Gen Psychiatry*. 2008; 65:329–336. [PubMed: 18316679]
- Deater-Deckard K, Petrill SA. Parent-child mutuality and child behavior problems: an investigation of gene-environment processes. *J Child Psychol Psychiatr*. 1999; 45:1171–1179. [PubMed: 15257673]
- Deater-Deckard K, Plomin R. An adoption study of the etiology of teacher and parent reports of externalizing behavior problems in middle childhood. *Child Dev*. 1999; 70:144–154. [PubMed: 10191519]
- Deater-Deckard, KM.; Pylas, MV.; Petrill, SA. Twin parent-child interaction system. Institute of Psychiatry; London: 1997.
- Deater-Deckard KM, Reiss D, Hetherington EM, Plomin R. Dimensions and disorders of adolescent adjustment: a quantitative genetic analysis of unselected samples and selected extremes. *J Child Psychol Psychiatr*. 1997b; 38:515–525. [PubMed: 9255695]
- Dick DM, Viken RJ, Kaprio J, Pulkkinen L, Rose RJ. Understanding the covariation among childhood externalizing symptoms: genetic and environmental influences on conduct disorder, attention deficit hyperactivity disorder, and oppositional defiant disorder symptoms. *J Ab Child Psychol*. 2005; 33:219–229.
- Du Fort GG, Boothroyd LJ, Bland RC, Newman SC, Kakuma R. Spouse similarity for antisocial behaviour in the general population. *Psychol Med*. 2002; 32(8):1407–1416. [PubMed: 12455939]
- Fagot BI, Leve LD. Teacher ratings of externalizing behavior at school entry for boys and girls: Similar early predictors and different correlates. *J Child Psychol Psychiatr*. 1998; 39:555–566. [PubMed: 9599783]
- Gardner F. Methodological issues in the direct observation of parent-child interaction: do observational findings reflect the natural behavior or participants? *Clin Child Fam Psychol Rev*. 2000; 3:185–198. [PubMed: 11225753]
- Goldsmith HH. A zygosity questionnaire for young twins: a research note. *Behav Genet*. 1991; 21:257–269. [PubMed: 1863259]

- Henderson SH. Method. *Monogr Soc Res Child Dev.* 1999; 64:26–49. [PubMed: 10685436]
- Hetherington EM, Clingempeel WG. Coping with marital transitions: a family systems perspective. *Monogr Soc Res Child Dev.* 1992; 57(2–3):1–242.
- Klump L, Burt SA. The Michigan State University Twin Registry (MSUTR): genetic, environmental, and neurobiological influences on behavior across development. *Tw Res Hum Genet.* 2006; 9:971–977.
- Leve LD. Observation of externalizing behavior during a twin-friend discussion task. *Mar Fam Rev.* 2001; 33:225–250.
- Leve LD, Winebarger AA, Fagot BI, Reid JB, Goldsmith HH. Environmental and genetic variance in children's observed and reported maladaptive behavior. *Child Dev.* 1998; 69:1286–1298. [PubMed: 9839416]
- McGue M, Bouchard TJ Jr. Adjustment of twin data for the effects of age and sex. *Behav Genet.* 1984; 14:325–343. [PubMed: 6542356]
- Miles DR, Carey G. Genetic and environmental architecture of human aggression. *J Pers Soc Psychol.* 1997; 7:207–217. [PubMed: 9008382]
- Neale, MC. *Mx: Statistical Modeling* [computer software]. Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University; Richmond: 1999.
- Neiderhiser JM, Reiss D, Hetherington EM. The Nonshared Environment in Adolescent Development (NEAD) Project: a longitudinal family study of twins and siblings from adolescence to young adulthood. *Tw Res Hum Genet.* 2007; 10:74–83.
- Nichols RC, Bilbro WC Jr. The diagnosis of twin zygosity. *Acta Genetica.* 1966; 16:265–275.
- O'Connor TG, Hetherington EM, Reiss D, Plomin R. A twin-sibling study of observed parent–adolescent interactions. *Child Dev.* 1995; 66:812–829. [PubMed: 7789203]
- Patterson GR, Forgatch MS. Predicting future clinical adjustment from treatment outcome and process variables. *Psychol Assess.* 1995; 7:275–285.
- Peeters H, van Gestel S, Vlietinck R, Derom C, Derom R. Validation of a telephone zygosity questionnaire in twins of known zygosity. *Behav Genet.* 1998; 28:159–163. [PubMed: 9670591]
- Pike A, McGuire S, Hetherington EM, Reiss D, Plomin R. Family environment and adolescent depressive symptoms and antisocial behavior: a multivariate genetic analysis. *Dev Psychol.* 1996; 32:590–603.
- Plomin R, Foch TT, Rowe DC. Bobo clown aggression in childhood: environment, not genes. *J Res Pers.* 1981; 15:331–342.
- Plomin, R.; DeFries, JC.; McClearn, GE. *Behavioral genetics: a primer.* W. H. Freeman; New York: 1990.
- Reiss, D.; Neiderhiser, JM.; Hetherington, EM.; Plomin, R. *The relationship code: deciphering genetic and social influences on adolescent development.* Harvard University Press; Cambridge: 2000.
- Rende RD, Slomkowski CL, Stocker C, Fulker DW, Plomin R. Genetic and environmental influences on maternal and sibling interaction in middle childhood: a sibling adoption study. *Dev Psychol.* 1992; 28:484–490.
- Rhee SH, Waldman ID. Genetic and environmental influences on antisocial behavior: a meta-analysis of twin and adoption studies. *Psychol Bull.* 2002; 128:490–529. [PubMed: 12002699]
- Rowe, DC.; Woulbroun, EJ.; Gulley, BL. Peers and friends as nonshared environmental influences. In: Hetherington, EM.; Reiss, D.; Plomin, R., editors. *Separate social worlds of siblings: the impact of nonshared environment on development.* Erlbaum; Hillsdale: 1994. p. 159–173.
- Rutter M, Macdonald H, Le Couteur A, Harrington R, Bolton P, Bailey A. Genetic factors in child psychiatric disorders-II. Empirical findings. *J Child Psychol Psychiatry.* 1990; 31:39–83. [PubMed: 2179248]
- Rutter M, Silberg J, O'Connor TJ, Simonoff E. Genetics and child psychiatry: I Advances in quantitative and molecular genetics. *J Child Psychol Psychiatry.* 1999; 40:3–18. [PubMed: 10102724]
- Towers H, Spotts E, Neiderhiser JM, Plomin R, Hetherington EM, Reiss D. Genetic and environmental influences on teacher ratings of the child behavior checklist. *Int J Behav Dev.* 2000; 24:373–381.

- Turkheimer E, Waldron M. Nonshared environment: a theoretical, methodological, and quantitative review. *Psychol Bull.* 2000; 126:78–108. [PubMed: 10668351]
- van den Oord EJ, Boomsma DI, Verhulst FC. A study of problem behaviors in 10- to 15-year-old biologically related and unrelated international adoptees. *Behav Genet.* 1994; 24:193–205. [PubMed: 7945150]
- van der Valk JC, Verhulst FC, Neale MC, Boomsma DI. Longitudinal genetic analysis of problem behaviors in biologically related and unrelated adoptees. *Behav Genet.* 1998; 28:365–380. [PubMed: 9926618]
- van Dulmen MHM, Egeland B. Analyzing multiple informant data on child and adolescent behavior problems: Predictive validity and comparison of aggregation procedures. *Int J Beh Dev.* 2011; 35:84–92.

Table 1

Demographic characteristics of the MSU, OTP, and NEAD samples

	MSU	OTP	NEAD
Percentage Caucasian	84%	92%	95%
Parent age in years	38.8 (5.4)	37.5 (5.0)	41.8 (6.2)
Mother education completed in years	13.0 (1.9)	14.5 (2.3)	13.6 (2.4)
Father education completed in years	12.8 (2.2)	15.0 (2.9)	14.0 (2.7)
Percentage single-parent families	19.5%	14%	0%

Note: Unless otherwise noted, scores are given as $M(SD)$

Table 2

Twin and sibling intraclass correlations for externalizing behavior by sample

Sample	MZ	DZ	FI	FS	HS	SS
MSU	0.42	0.44	-	-	-	-
<i>N</i>	41	59				
OTP	0.73	0.80	-	-	-	-
<i>N</i>	77	72				
NEAD	0.66	0.56	0.47	0.43	0.58	0.39
<i>N</i>	92	94	90	173	105	124

MZ Monozygotic twins, *DZ* dizygotic twins, *FI* full siblings in nondivorced families, *FS* full siblings in stepfamilies, *HS* half siblings, *SS* step siblings

Table 3

Parameter estimates for genetic, shared environmental, and nonshared environmental influences on observed externalizing behavior

Sample	A	C	E
<i>Unconstrained</i>			
MSU			
Full Model	0.00	0.44	0.56
	[-0.52, 0.52]	[-0.58, 0.58]	[0.38–0.74]
Best-fitting Model	–	0.44	0.56
	–	[0.27, 0.58]	[0.42, 0.73]
<i>OTP</i>			
Full Model	0.00	0.77	0.23
	[-0.17, 0.17]	[0.61, 0.83]	[0.17, 0.30]
Best-fitting Model	–	0.77	0.23
	–	[.69, .83]	[0.17, 0.31]
<i>NEAD</i>			
Full Model ^a	0.31	0.38	0.30
	[.14, .46]	[.28, .48]	[0.22, 0.41]
<i>Constrained</i>			
All samples	0.25	0.42	0.33
	[0.12, 0.36]	[0.34, 0.50]	[0.27, .040]

^aThe full model was the best-fitting model for the NEAD sample. Confidence intervals are provided in brackets