Parenting and child cognitive and socioemotional development:

A longitudinal twin differences study

Cassandra M.D. Hart

Human Development and Social Policy
Northwestern University

Working draft
September 19, 2008

Abstract

This paper uses data from the Early Childhood Longitudinal Study-Birth Cohort to analyze the relationship between parenting directed toward 9-month-old children and their subsequent cognitive and socioemotional development at 24 months. This relationship is tested using multiple regression on a nationally representative sample of children (n=9832), and on a subsample of twins (n=1520, 760 pairs). The twin subsample is also used to test the relationship using a twin differences model. Parenting practices are not significantly associated with socioemotional functioning at 2 years once child development at 9 months is controlled. OLS models yield a significant relationship between parenting and cognitive outcomes that diminishes as additional controls are included. This relationship disappears almost entirely when twin differences are employed, suggesting that the OLS findings result partly from a failure to control for unobserved family characteristics. Four potential moderators—race, SES, child temperament, and gender—are tested; none produces consistent evidence of moderation.

Parenting practices directed towards children are important in fostering child development (Landry, Smith, MillerLoncar, & Swank, 1997; Wacharasin, Barnard, & Spieker, 2003). Parenting quality is especially important for infants, because they spend more time in parental care than do older children (Bornstein, 2002). This paper analyzes the relationship between parenting behaviors in children's first year of life and their subsequent cognitive and socioemotional development. This study builds upon past research by using a twin sample to eliminate omitted variable bias due to unobserved family characteristics.

Parenting and cognitive development

The quality of parent-child interactions is related to cognitive outcomes both cross-sectionally and longitudinally (Elardo, Bradley, & Caldwell, 1977; Kelly & Barnard, 2000; Kelly, Morisset, Barnard, Hammond, & Booth, 1996; Wacharasin et al., 2003). For instance, maternal supportiveness longitudinally predicts children's performance on cognitive development tests (Mistry, Biesanz, Chien, Howes, & Benner, 2008; Wacharasin et al., 2003). Parents' teaching behaviors are also associated with IQ in later childhood (Bee et al., 1982; Marc H. Bornstein, 1985; Kelly & Barnard, 2000).

Parenting is associated with specific academic skills as well. Parental responsiveness and learning stimulation are associated with children's language skills (Kelly & Barnard, 2000; Tamis-LeMonda, Bornstein, & Baumwell, 2001) and scores on vocabulary, reading, and math tests (Bradley, Corwyn, Burchinal, McAdoo, & Coll, 2001). Maternal teaching behaviors, such as scaffolding child learning and directing children's attention during play, are also related to children's early attention skills (Allhusen et al., 2003; Findji, 1993; but see Wakschlag & Hans, 1999). Attention skills at school entry are an important predictor of school readiness and subsequent academic achievement (Allhusen et al., 2003; Duncan et al., 2007).

Parenting and socioemotional and behavioral development

Parenting has also been linked to socioemotional and behavioral outcomes for young children (Wacharasin, 2001). Harsh or negative parenting is associated with children's behavior problems, including anxiety, less prosocial behavior, hyperactivity, and conduct problems (Asbury, Dunn, & Nomin, 2006; Asbury, Dunn, Pike, & Plomin, 2003; Deater-Deckard et al., 2001; Pike, Iervolino, Eley, Price, & Plomin, 2006). Restrictive parenting also negatively affects children's social development, such as social initiation skills (Landry et al., 1997).

Parenting need not be overtly harsh or negative to impair children's socioemotional development. Less positive parenting climates (such as less parental warmth, less responsive parenting, less positive parental feelings towards the child, or less positive communication strategies) are associated with children's socioemotional functioning, including disruptive behavior, positive or negative temperament, prosocial behavior, noncompliance, and social skills (Asbury et al., 2006; Deater-Deckard et al., 2001; Pike et al., 2006; Wakschlag & Hans, 1999). These relationships have been found both cross-sectionally (Asbury et al., 2003; Deater-Deckard et al., 2001), and longitudinally (Asbury et al., 2006; Landry et al., 1997; Pike et al., 2006).

Potential moderators

Other factors may moderate the relationship between parenting and child outcomes. Culture (proxied by race or ethnicity) may affect the meanings that children attach to parenting practices, and therefore the developmental sequelae of those practices (Ispa et al., 2004). For instance, physical discipline is associated with negative behavioral outcomes for White children, but not for Black children (Deater-Deckard, Bates, Dodge, & Pettit, 1996). Home environment measures, including parenting, are differentially predictive of cognitive outcomes for white, Black, and Mexican-American families as well (Bradley, Caldwell, & Rock, 1988).

Socioeconomic status may also moderate the relationship between parenting and child development. Poverty introduces a host of risks that militate against the full realization of children's potential (Chase-Lansdale, Wakschlag, & Brooks-Gunn, 1995). The effects of poor parenting may be magnified by these other risks, and poor children may therefore suffer more from sub-optimal parenting than do children in high-SES families. Alternatively, the stresses of poverty may so depress children's potential that even high quality parenting cannot counteract these stressors enough to ensure that children reach their full potential. In this case, parenting quality may be more strongly related to child outcomes in higher-SES households. Past research has offered some support for the latter hypothesis (Bradley et al., 1988).

A third potential moderator is the temperament of the focal child. The quality of parent-child interaction is affected by the personal characteristics that both the parent and the child bring to the interaction, and research suggests that child characteristics affect which parenting practices best support child development (Collins, Maccoby, Steinberg, Hetherington, & Bornstein, 2000; Kochanska, Friesenborg, Lange, & Martel, 2004).

Finally, gender may moderate the relationship between parenting and child outcomes. Research suggests that girls are more responsive to sensitive parenting than are boys (Bradley et al., 1988), and that parents treat boys and girls somewhat differently. For instance, parents are more likely to verbally engage their daughters than their sons (Gleason & Ely, 2002; Maccoby, 2003). Given this tendency, different meanings may be attributed to parenting behaviors based on child gender. Examining differences in effects by gender is therefore important.

Omitted variable bias

As reviewed above, most research indicates that positive parenting practices are linked to improved cognitive and socioemotional outcomes, concurrently and longitudinally. However, many of these studies are vulnerable to omitted variable bias; unmeasured characteristics associated with both parenting practices and the outcomes of interest may drive the associations found in research. Moreover, studies using cross-sectional designs are plagued by the possibility of reverse causality: differences in parenting practices may arise *in response* to differences in children's qualities, rather than causing those qualities (Reiss, Neiderhiser, Hetherington, & Plomin, 2000). Recently, researchers have tried to address these problems by using sophisticated methods like sibling fixed effects designs or longitudinal analysis (Asbury et al., 2006; Caspi et al., 2004). Some studies also minimize single-source bias by using multiple reporters to establish children's developmental trajectories (Asbury et al., 2006; Caspi et al., 2004)

Following this line of well-controlled work, this study tests whether differential parental treatment of children predicts differential child development. Specifically, I hypothesize that positive parenting behavior predicts positive cognitive and socioemotional development, and that this relationship will help explain within-family differences in development in a twin sample. That is, I predict that the twin who receives more favorable treatment at 9 months will have an advantage in cognitive and socioemotional development relative to the less favored sibling.

This study builds upon past research in several ways. First, it uses twin data to eliminate bias due to unmeasured family characteristics that are associated with both outcomes and predictor variables. Second, it employs longitudinal analysis, obviating the possibility that outcome measures could be affecting predictor measures rather than vice versa.

Finally, it measures parenting behavior towards children at a younger age (9 months) than have prior articles using both twin samples and longitudinal methods (e.g., Asbury et al., 2006; Burt, McGue, Iacono, & Krueger, 2006; Caspi et al., 2004; Liang & Eley, 2005). To the greatest extent possible, this mitigates the possibility that differential parental treatment of children is evoked by differences in children's personalities or skills caused by child-specific environmental influences (such as illness or accidents), because taking measurements at this early age minimizes the time that children may have been exposed to different environments. Even at 9 months, there is some divergence in twins' experiences, but less so than if data were collected after several years of differential experiences had accrued.

A disadvantage of this study compared to some prior studies is that I use both monozygotic and dizygotic twins, rather than restricting my sample only to monozygotic twins. This was done to retain a sufficiently large sample; the ECLS-B includes less than 250 monozygotic twin pairs. I tested an interaction in the twin differences model to determine whether zygosity moderated the relationship between differences in parenting and differences in the outcomes. A significant interaction would indicate that differences in parenting quality matter differently based on whether the twins are genetically identical. However, since this interaction was not significant, I combined monozygotic and dizygotic twins for the full analysis.

It is important to note that while this study uses twins as a sample population, it is not a genetic study. It does not exploit information about the degree of relatedness of various subjects. Rather, it is a *sibling differences* study, with a sibling population restricted to twins. This may offer an advantage over traditional sibling differences studies because twins, whether monozygotic or dizygotic, experience shared family influences at the same time.

For instance, research suggests that family income in early childhood is more predictive of children's educational attainment than is income in later childhood (Duncan, Yeung, Brooks-Gunn, & Smith, 1998). Unlike siblings born years apart, twins will experience the same levels of family income *at the same time*. Although twins do not necessarily experience environments in exactly the same way, even when their environments are objectively the same (Turkheimer &

Waldron, 2000), using twins establishes a greater similarity between objective environments than is experienced by siblings who are exposed to the same environmental influences at different points in their development. Thus, while this study is not genetically informed, it provides controls superior to OLS models, and possibly superior to traditional sibling difference models.

DATA

The study sample is drawn from the Early Childhood Longitudinal Study-Birth Cohort (ECLS-B). The ECLS-B follows a cohort of 10,688 children born in 2001 from their birth until kindergarten. When sample weights are employed, the sample is nationally representative of children born in 2001, excluding children who died before 9 months, who were placed for adoption, or who were born to mothers younger than 15 years old.

Data collection occurred at 9 months, 2 years, preschool age (4 years) and kindergarten age (5 years). Information was also drawn from children's birth certificates. This analysis relies on data drawn from birth certificates, as well as the 9-month and 2-year waves of data collection.

Analytic samples

This study uses two samples. The first includes all children in the ECLS-B who contributed data at both waves of data collection, including twins and singleton births (n=9,832). The second consists of only the twins sampled in the ECLS-B, including both monozygotic and dizygotic twins (n=1520; 760 pairs). The ECLS-B oversampled twin births to allow analysis on this group to be carried out with precision (Andreassen & Fletcher, 2005).

Dependent measures

This study examines two outcomes. Cognitive development is measured using the mental subscale of the Bayley Short Form-Research Edition (BSF-R) scale. The BSF-R was adapted from the Bayley Scales of Infant Development, 2nd Edition (Andreassen & Fletcher, 2007). The mental index measures performance on a number of domains, including jabbering expressively, early problem solving, and receptive vocabulary (Andreassen & Fletcher, 2007). This score was standardized using the standard deviation found for the full sample.

The mental index demonstrated high internal consistency for the full ECLS-B sample (r_{xx}=.88) (Andreassen & Fletcher, 2007). While lack of access to individual items prevented me from calculating the reliability using item-level correlations for the twin subsample, using subscale scores as items established high reliability for the Bayley Mental scale in the twin subsample (Cronbach's α=.89).

Child socioemotional functioning is measured using interviewer ratings of children's behaviors during administration of the Bayley Mental and Motor Development tests. Interviewers rated the extent to which children displayed each of eleven behaviors. Sample behaviors include positive and negative affect, attention to tasks, persistence, social engagement, and frustration (Andreassen & Fletcher, 2007). Scores on these items are added to create a composite, the Behavior Rating Scale (BRS). Higher scores indicate better socioemotional functioning. This score was also standardized using the standard deviation found in the full sample. While NCES did not report the reliability of this scale for the full sample, analysis using the twin subsample indicates that the composite forms a coherent scale (α=.91).

Independent measures

The key independent measure is the parenting subscale of the Nursing Child Assessment Teaching Scale (NCATS). The NCATS measures parent-child interaction during a teaching task (Barnard et al., 1989; Kelly & Barnard, 2000). Parents selected a task that the child could not yet perform from a list of 15 potential teaching activities, and taught the activity to the child

(Andreassen & Fletcher, 2005). The teaching task was videotaped, and trained coders rated whether dyads exhibited each of 73 behaviors. The parent subscale consisted of 50 items relating to parents' teaching behaviors and responsiveness (Andreassen & Fletcher, 2005; Barnard et al., 1989). The NCATS had good inter-rater reliability; coders achieved at least 85 percent agreement on weekly reliability checks (Andreassen & Fletcher, 2005). However, the reliability of the parenting subscale was rather low (Cronbach's α =.68) (Andreassen & Fletcher, 2005). An alpha of .7 is generally considered acceptable (Nunnally & Bernstein, 1994). When measured in the twin subsample, the scale was marginally above the .7 threshold (α =.72).

Parent NCATS scores predict child IQ as late as 5 years of age, and differentiate between mothers on characteristics like education, marital status, substance abuse problems, and family stress (Barnard et al., 1989; Kelly & Barnard, 2000). The scale's ability to predict socioemotional outcomes has been less tested. However, other studies have found associations between constructs captured by the teaching scale (e.g., responsiveness to distress, fostering social-emotional growth, sensitivity to child's cues) and children's behavioral and socio-emotional outcomes (Barnard et al., 1989; Landry et al., 1997; Wakschlag & Hans, 1999). Thus, I posit that parent NCATS scores will predict both cognitive and socioemotional child outcomes.

Control measures-demographics

Standard demographic controls were employed. For OLS regressions, controls included child birth weight, child sex, child age, race, family income, parents' marital status, maternal age at the child's birth, and maternal education at the child's birth. Child sex (male=1), birth weight (in pounds), years of maternal education, and maternal age were obtained from birth certificate data. Maternal education is top-coded at 17 years of education. Marital status (married=1) and parental income were obtained through parent reports. Parental income is reported in increments of $5-10,000, and I assigned respondents to the midpoint of the range. Income was top-coded at $200,000. Income was divided by $10,000 for ease of interpretation of results. Child race was determined based on the birth certificate reports of parents' ethnic designations (Andreassen & Fletcher, 2005); dummy variables indicated whether the child was black, Hispanic, or other. Non-Hispanic whites are the omitted category.

For twin difference models, only child sex, birth weight, and child age were included. There is no between-twin variability in the other demographic characteristics (e.g., race, maternal age, or income), so any differences in twins' development cannot be explained by these family measures using a twin difference model. In the vast majority of cases, twins were tested on the same day, and were therefore exactly the same age when tested. However, 57 twin pairs were tested on different days. Only 14 pairs were tested more than 10 days apart.

Control measures-9-month child development variables

Lagged versions of the outcome variables, as well as other child development measures taken at 9 months, were also included as predictors in certain models. These measures control for child-specific heterogeneity in skill development. Controlling for these measures ensures that the effect of parenting at nine months is measured independent of the child's endowments and other influences that have promoted differential achievement up to that point. Because these lagged control measures are scales, they were all standardized using their respective standard deviations found for the full sample.

Nine-month Bayley Mental scores were gathered in the same way as in the two-year data collection wave (Andreassen & Fletcher, 2007). The reliability of the 9-month Bayley Mental scale in the full sample was r_{xx}=.81 (Andreassen & Fletcher, 2007). The scale exhibited lower reliability in the twin sample (α=.71).

Interviewer ratings on the Behavior Rating Scale were gathered at the 9-month data collection. The process of data collection and most items were the same as at the 2-year data collection. A few behaviors scored at the 2-year data collection were not scored at 9 months. These measures were fearfulness, frustration, persistence, and cooperation (Andreassen & Fletcher, 2007). Reliability for the 9-month BRS was acceptable for the full ECLS-B sample (α=.79) (Andreassen & Fletcher, 2005) and for the twin sample (α=.78).

The third 9-month child development control measure was the Infant/Toddler Symptom Checklist (ITSC), which measures child temperament and self-regulatory skills. Parents reported how characteristic certain behaviors were for their children; sample behaviors include fussiness, distractibility, and inability to wait for food or toys (Andreassen & Fletcher, 2005). Parents rated their children on a scale ranging from "never" (0) to "is like this most times" (3). Scores were summed to obtain a total rating. Higher scores reflect a more difficult temperament. The scale had rather low reliability in both the full sample (α =.63) and the twin subsample (α =.69).

Two additional child assessments were included as controls. Motor skill development at 9 months was measured using the BSF-R Motor subscale. The reliability was high scale for both the full sample (r_{xx}=.94) (Andreassen & Fletcher, 2005), and for the twin subsample (α =.87). The child subscale of the NCATS was also included. This scale had low reliability in the full sample (α =.62) (Andreassen & Fletcher, 2005) and the twin subsample (α =.64).

Analytic plan

For each outcome, two sets of populations are used to generate a total of six models. The first population used is the full sample for the ECLS-B; the second is the ECLS-B subsample of twins. In each table, the first two models presented are bivariate OLS regressions testing the relationship between parenting scores and the outcome for each of these populations, respectively. The third model adds demographic controls to the OLS regression for the full sample. While models with demographic controls were also tested for the twin samples, they are not shown here for ease of presentation. The fourth and fifth models include the parenting predictor, demographic controls, and the 9-month child development controls for the full and twin samples, respectively. I refer to these last two models as "fully-controlled OLS models".

The final regression uses a twin difference model. The twin difference estimation employs first differencing of twins' scores to control for family fixed effects. Difference scores were constructed by designating one twin as Twin 1 and the other as Twin 2, and subtracting Twin 2's score from Twin 1's score. These differences were then used as the dependent variable and independent variables in the model. I suppressed the constant in the twin difference models.

For all analyses, jackknife standard errors were used to adjust for the fact that the ECLS-B does not use a simple random sample. All analyses employed sampling weights. In OLS regressions, appropriate child-level weights were used; twin difference models used the mean of the weights assigned to each twin. Missing value dummies were employed to preserve information from cases with missing values on control measures. Cases with missing values on the NCATS parenting score for either twin were excluded from analyses.

RESULTS

Descriptive statistics are presented in table 1. Means and standard deviations are reported for both the full ECLS-B sample (including singletons as well as twins) and for the twin subsample. T-tests were conducted to determine whether the means were comparable for the twin subsample and the singletons in the full sample. Compared with singletons in the full

ECLS-B sample, twins had significantly lower cognitive scores and poorer socioemotional functioning ($p<.01$ for both). Parenting scores did not differ significantly between the groups.

There were also many significant differences among the set of child-specific controls. Twins scored significantly lower on the 9-month Bayley Mental and Motor tests, on the 9-month Behavior Rating Scale, and on the child portion of the NCATS ($p<.01$ for all measures). Twins and singletons did not have significantly different scores on the Infant/Toddler Symptom Checklist. Twins had significantly lower birth weight than did singletons, and they were also significantly younger when tested at wave 2 ($p<.01$ for both). There were no significant differences between twins and singletons in the sex compositions of the samples.

Family characteristics also differed in many cases. The twin sample had a significantly greater proportion of whites ($p<.01$), and lower proportions of Hispanics ($p<.01$) and Asians ($p<.05$) than did singletons. Significantly more parents in the twin sample were married ($p<.01$). Family incomes, maternal education, and maternal age were also higher for twins ($p<.01$).

While many significant differences were found, the size of the sample should be considered. Even the measure with the most missing data still had data for 1,224 twins and 6,788 singletons. Because the samples are so large, the standard errors are very small, and relatively small mean differences between twins and singletons can be detected with precision.

Establishing reliability for twin difference measures

One major concern when sibling difference models are used is establishing acceptable reliability. Measurement error in independent variables is of particular importance. Measurement error in dependent variables does not bias coefficients; it merely inflates standard errors (Bound, Brown, & Mathiowetz, 2001). However, measurement error in explanatory variables does result in biased estimates. When measurement error is unrelated to the “true”, unobserved value of the explanatory variable, regression estimates are biased downward (Bound et al., 2001; Wooldridge, 2003). The reliability coefficient of an explanatory variable represents the factor by which an estimate would be attenuated either in a bivariate regression, or in a multivariate regression where that explanatory variable was uncorrelated with all other predictors (Bound et al., 2001). In this simple case, the coefficient of an explanatory variable with reliability r_{xx}=.7 should be inflated by $\frac{1}{.7}$ to determine the coefficient's true magnitude when measured without error. The attenuation effect is more complicated in cases where the explanatory variables are not uncorrelated (Bound et al., 2001).

Problems with mismeasurement may be especially acute in first differences analyses because difference scores are generally less reliable than their constituent variables (Allison, 1990; Bound et al., 2001; Cronbach & Furby, 1970). Therefore, it was important to establish whether the difference measures used in this analysis were sufficiently reliable to generate credible results. Scales are used to measure both child outcomes, the NCATS parenting score at 9 months, and the 9-month child development control variables. Because the ECLS-B provides either the constituent items or subscales on all of these measures, I was able to use differences in the constituent items to establish an alpha for the *difference score* for the full scale.

This analysis indicated that a reasonable degree of reliability exists for the difference scores used in the regressions. According to Nunnally and Bernstein (1994), a reliability score of .7 is considered acceptable, although lower reliabilities (e.g., alpha=.5 or .6) may be accepted for early research on a given construct. Table 2 presents Cronbach's alphas for dependent and independent measures for the full ECLS-B sample (as measured by the NCES team of psychometricians), the twin subsample used in OLS regressions, and the twin difference measures. Difference score reliabilities are very high for two of the child outcome measures (alphas greater than .85 for the Bayley Mental Scale and the Behavior Rating Scale). Reliability for the parenting score on the NCATS also falls only slightly short of the .7 threshold (α=.65). Two of the five control scales—the 9-month interviewer behavior ratings and the 9-month Bayley Motor score—have reliabilities above .7, while the Mental score (α=.62), the Infant/Toddler Symptom Checklist score (α=.58), and the NCATS child score (α=.56) fall short of the normal threshold for standard research, but above the recommended threshold for exploratory research.

Establishing variability for twin differences

I next established that sufficient differences exist between twins to conduct the first differences analyses. If there were very little variation in how parents treated their twins, or in twins' cognitive and socioemotional functioning scores, it would be difficult to use twin differences to detect relationships between parenting and child outcomes. However, an appreciable amount of variation between twins does indeed exist. Figures 1-3 shows the distribution of absolute differences between twins on the outcomes (Bayley Mental Scale score and Behavior Rating Scale at 2 years) and the key independent variable (NCATS parenting score at 9 months), given in terms of standard deviations of the full group. In regression analyses, the direction of the difference between twins is important to establish whether advantages in parenting are associated with advantages in later performance, so the relative difference between twins is used in that part of the analysis. However, for the purpose of establishing the *degree* of difference between twins, the absolute value of the differences in a twin pair was used.

On average, there was a gap between twins' Bayley Mental Scale scores of 6.14 points. This is equivalent to .58 of the standard deviation for Bayley Mental Scale (10.65 points) found for the full sample (see Figure 1). Twins' Behavior Rating Scale scores were even more dissimilar (see Figure 2); twins had, on average, a .7 standard deviation difference in this score (5.81 raw points). Twins had the biggest difference in the independent measure. NCATS parenting scores were, on average, .85 standard deviations (3.80 raw points) apart for twins. The fact that twins did experience differences in both parenting and the outcome variables suggests that it is reasonable to use twin differences to estimate the effect of parenting.

Correlations

Correlations are presented in table 3. Below the diagonal, correlations between measures are reported for the full sample. Correlations between *twin differences* of the same measures are reported above the diagonal. Because twins do not vary on demographic characteristics, correlations between twin differences are omitted for this set of variables. Correlations are smaller in magnitude when twin differences are used. Furthermore, far fewer of the twin difference correlations are significant. For instance, the 2-year Bayley Mental measure is significantly correlated with all predictor variables in the full sample. When twin differences are used, it is only significantly related to birth weight, child sex, and 9-month motor skills. Notably, neither outcome variable is significantly correlated with the 9-month NCATS parenting score when twin differences are used. Both correlations are also very small in a substantive sense (r=.00 for the Bayley Mental Scale; r=.01 for the BRS).

Regressions

OLS regressions indicate that the NCATS parenting score positively predicts Bayley Mental scores at 2 years for the full sample and the twin sample, at all levels of controls (Table 4, Columns 1-5). While the magnitude of the coefficient on the NCATS parenting score drops as each successive level of controls is added, the relationship remains significant through the fully controlled OLS models (Columns 4 and 5).

When twin differences estimation is used to account for unmeasured family fixed effects, the relationship between differences in parenting and differences in mental development found in these OLS equations disappears (Table 4, Column 6). The coefficient on the NCATS parenting scale (b=.004) was both statistically insignificant and substantially smaller than when OLS analysis was used. Parenting scores and the outcome variables are standardized by the full sample standard deviation. Thus, this coefficient means that a one standard deviation advantage in parenting for Twin 1 compared to Twin 2 translates into only a .004 standard deviation advantage for Twin 1 in the 24-month cognitive measure, on average.

The failure to find a significant effect is not due simply to a lack of power. Because both the parenting variable and the outcome variables are standardized, the standard errors can be used to determine the amount of predictive power afforded by the model. The model has the power to detect significant differences when coefficients are approximately double the size of the standard error. For instance, the standard error on the parenting coefficient in the twin differences model is .047, so a significant difference would be detected if the coefficient on parenting were approximately .1 or higher. In standardized terms, an effect size of .1 is rather small; thus, the twin differences model gives us power to detect even differences that are considered small by standard guidelines (Cohen, 1992). The failure to find an effect is more notable in light of the power provided by this model.

All child-specific demographic variables significantly predict 2-year mental scores in all models. Child sex, birth weight, and age remain significant predictors of Bayley Mental scores when the more stringent first difference model is used. Moreover, the magnitude of the coefficient on birth weight actually *increases* when first differences are used. These findings are consistent with past research (e.g., Breslau, Chilcoat, DelDotto, Andreski, & Brown, 1996;

Galsworthy, Dionne, Dale, & Plomin, 2000). This adds further credibility to the null result for the parenting variable, because it rules out the possibility that the null finding is due to the twin difference model's inability to detect *any* significant relationships predicted by past research.

Among 9-month child scores, the mental score is a significant predictor in both the full-sample and twin-subsample OLS models, but not in the first difference model. The motor score, which is significant in neither the full-sample nor the twin-sample OLS models, becomes significant in the twin differences model (b=0.129, $p<.05$). Nine-month BRS, NCATS child, and ITSC scores significantly predict mental scores in the full-sample OLS model only.

Family controls, which could not be tested by the twin differences model, yielded a few consistently significant predictors of mental scores at 2 years. Family income and maternal education both remain statistically significant, positive predictors of cognitive scores in the fully controlled full-sample and twin-subsample OLS models (Columns 4 and 5). Holding all else equal, African-Americans and Hispanics in the full sample both score significantly lower than their non-black, non-Hispanic counterparts. For the twin subsample, the coefficient on Hispanic drops to marginal significance, while the coefficient on Black remains significant.

The other outcome tested is child socioemotional functioning. For both the full sample (Table 5, Column 1) and the twin subsample (Table 5, Column 2), bivariate regressions indicate that parenting at 9 months is significantly associated with BRS scores at 2 years. While this relationship is somewhat attenuated when demographic controls are added (Column 3), it remains statistically significant. However, when the 9-month child development measures are added, the relationship dwindles to marginal significance for the full sample (Column 4), and non-significance for the twin sample (Column 5). When twin differences are employed (Column 6), the relationship between parenting and subsequent behavior remains insignificant.

In all models, male sex is associated with poorer socioemotional functioning at 2 years. Child age, maternal education, and the child's 9-month BRS scores are significant predictors in the full sample and twin OLS models with full controls (Columns 4 and 5). Birth weight, income, and married parents are significant, positive predictors of socioemotional skills in the full sample, fully controlled OLS model, while the 9-month rating of self-regulation problems (ITSC) negatively predicts behavior at 2 years (Column 4). These predictors dropped to non-significance or marginal significance in the twin sample.

Interactions

A series of interactions was tested to determine whether race, SES, child temperament, and gender moderate the relationship between parenting and subsequent child outcomes.

Race: To test whether the relationship between parents' teaching behaviors and cognitive and socioemotional development differs by race/ethnicity, two separate interaction models were run. One included an NCATS*Black interaction term; the other used an NCATS*Hispanic interaction term. Since twins share an ethnic designation, the main effects of the race variables are not estimated in the twin differences models. Results are shown for the fully-controlled OLS models for the full and twin samples, and for the twin difference model. These models are the best controlled, so they present the most rigorous test of whether interactions actually exist.

For the cognitive outcome, the NCATS*Black interaction term is not significant in any of the well-controlled models (Table 6), although it is marginally significant ($p<.10$) in the OLS model for the twin sample. While the main effect of parenting is significant in the OLS models for both the twin and full samples, the coefficient drops to .000 in twin differences models.

The results are even less indicative of an interaction for African-American vs. other families when socioemotional functioning is used as an outcome. The interactions are not even marginally significant in any of the well-controlled models (results available upon request). Indeed, even in the OLS models that lack lagged child measures or demographic controls, the interaction term is never more than marginally significant (results available upon request).

A second model tested the interaction of the NCATS parenting score with a dichotomous variable indicating whether the child was Hispanic vs. non-Hispanic. However, in no model—well-controlled or otherwise—was this interaction significant for either the 2-year cognitive or the socioemotional outcome measures (results available upon request).

These results suggest that the importance of parenting for child outcomes does not differ between Hispanic and non-Hispanic families, or between black and non-black families.

Socioeconomic status: To test whether the relationship between parenting and cognitive and socioemotional outcomes differed by socioeconomic status, a series of moderators were used. The first measured whether families were below the poverty line (Poverty=1 if families were below the poverty line). Approximately one quarter of families in both the full sample (23.18%) and the twin sample (24.31%) were below the poverty line. The second moderator measured whether families were below 130% of the poverty line. For both the full sample and the twin subsample, slightly over one-third of families were below 130% of the poverty line (34.72% for the full sample; 34.78% for the twin sample). The third moderator measured whether families were below 185% of the poverty line. Nearly half of the families in both the full sample (47.12%) and the twin sample (45.93%) had incomes below 185% of the poverty line. These cut-off points were chosen in line with ECLS-B-provided variables.

When the poverty line and 130% of the poverty line were used as the relevant thresholds, there were no significant interactions in any of the models, for either outcome (results available upon request). However, a threshold of 185% of the poverty line yielded a few significant interactions. For the Bayley Mental Scale, there was a significant negative interaction between parenting and being below 185% of the poverty line in the fully-controlled OLS model using the full sample (Table 7, Column 1). However, this interaction was not significant in either the twin OLS model or the twin differences model (Table 7, Columns 2 and 3). The significance did not decline solely due to inflated standard errors associated with smaller samples; the magnitudes of the coefficients also fell. Moreover, in the less-controlled versions of the full-sample OLS model (not shown), the interaction is either non-significant or only marginally significant. This suggests that this interaction may be a chance finding rather than a genuine result.

Interestingly, there is also a significant interaction in one model for the socioemotional functioning outcome when 185% of the poverty line is used as the threshold. While the interaction is not significant in any of the OLS models for either the full sample or the twin sample, the interaction becomes significant when twin differences estimation is used (Table 8, Column 3; $b = -.241$, $p < .01$). Furthermore, the main effect on the NCATS parenting variable becomes significant ($b = 0.129$, $p < .05$). This result indicates that receiving better parenting is significantly beneficial for a twin's socioemotional outcomes only in families with incomes *above* 185% of the poverty line. For families with incomes below 185% of the poverty, better parenting is actually negatively associated with socioemotional functioning.

However, there are several reasons to be cautious in accepting this result. First, the effects are not mirrored in the OLS models. A consistent pattern across the models would increase confidence in this result. Second, it is theoretically puzzling why a difference between high-SES and low-SES families should be absent when 100% of the poverty line and 130% of the poverty line are used as thresholds, and emerge only when the definition of low-SES is expanded to include families with incomes below 185% of the poverty line. Additional analyses using twin differences (results available upon request) confirm that between three categories of lower-income families—those below the poverty line, those with incomes between 100% and 130% of the poverty line, and those with incomes between 130% and 185% of the poverty line—the group with incomes between 130-185% of the poverty line had relationships between parenting and socioemotional development that *least* resembled those in the high income group. While parenting scores were significantly and positively related to BRS scores in the high income group, the relationship was significantly negative for families with incomes between 130-185% of the poverty line. This result is counter-intuitive, and should be researched further.

Child temperament/self-regulation: The moderating role of child temperament was tested using a dummy variable indicating whether the child's ITSC score was greater than the mean, suggesting below-average self-regulation skills (1=below-average skills). This term was interacted with parenting to determine whether parenting differentially affects outcomes based on child temperament. Slightly over one third of the twin pairs (34.33%) were discordant on the mean split variable, maintaining sufficient variability to conduct the twin differences analysis.

For the cognitive outcome, interactions between parenting and temperament were not significant in any of the fully-controlled OLS models, or for the twin differences model (results available upon request). Thus, temperament did not moderate the relationship between parenting at 9-months and subsequent cognitive development.

For the socioemotional outcome, the interaction terms were non-significant for the OLS models for both the full and twin samples. This was true in the fully-controlled OLS models (Table 9, Columns 1 and 2), as well as in the less-controlled OLS models (results available upon request). However, when twin differences estimation was used, the interaction term became significant ($b = -.155$, $p < .01$; Table 9, Column 3). This term indicates that if Twin 1 (but not

Twin 2) has below-average self-regulation skills, and Twin 1 receives better parenting than Twin 2, Twin 2 is predicted to have lower socioemotional scores at 2 years, beyond the (insignificant) decrement associated with having worse self-regulation at 9 months. By contrast, the main effect of parenting in the twin differences model suggests that if both twins have the same status for temperament problems, better parenting predicts (insignificantly) better socioemotional scores at 2 years. Figure 4 gives a visual illustration of this interaction. The “worse self-regulation” line represents children with below-average self-regulation skills, whose twins have above-average self-regulation skills. The “better self-regulation” line represents children with above-average self-regulation skills, whose twins have below-average self-regulation skills. The “even” line represents cases where both twins are above-average, or both are below-average.

To check the robustness of the interaction between parenting and child temperament for the socioemotional outcome, an alternative version of the interaction was estimated using the continuous ITSC score as the moderator. This specification yielded a marginally significant interaction in the twin differences model ($b= -.040$, $p<.10$; results available upon request). The interaction term for the fully-controlled OLS model using the full sample was also marginally significant, but the effect was in the opposite direction ($b= .024$, $p<.10$). These results do not provide strong support for a parenting-temperament interaction.

Child sex: Child sex was also tested as a moderator (results available upon request). However, the interaction term was not significant for any of the models, for either outcome.

DISCUSSION

This paper presents results from a nationally representative dataset examining the link between parenting behaviors when children are 9 months old and child cognitive and socioemotional development at 2 years. It finds that in the best controlled models, the relationship between parenting behavior and child outcomes disappears. While neither race nor gender moderated the effects of parenting behaviors on subsequent outcomes, there was some evidence that income measures and temperament measures may moderate these relationships. However, these findings were not particularly robust to alternative specifications of the moderating variables.

These results suggest that evidence linking parenting to child outcomes may suffer from omitted variable bias. Twin differences estimates eliminate omitted variable bias derived from unmeasured family characteristics common to both twins. When this estimation method is used, the relationship between parenting and child cognitive outcomes disappears.

The relationship between parenting and child socioemotional development disappears using even less stringently controlled models. This association was significant when demographic characteristics were controlled, but dropped to insignificance when lagged measures of child development were included. This suggests that studies that lack data on prior child development may find relationships between parenting and child socioemotional outcomes that would vanish if prior child measures were included.

Several objections may be raised to this study's conclusions. For instance, the use of a twin sample in the twin differences models may limit the generalizability of these results. Twins present unique challenges to parents; parenting processes may differ when parents are confronted with two children of the same age (Lytton & Gallagher, 2002). However, parenting scores for twins were not significantly different than scores for singletons, suggesting that parenting quality in the twin sample is not systematically different than for singletons. Moreover, the threat to external validity posed by using twin differences must be balanced against the improvement in internal validity: twin differences models provide a better test of whether parenting is causally related to cognitive and socioemotional outcomes than do OLS models.

Another possible objection is that measurement error may be partly responsible for the lack of relationship between parenting at 9 months and later child outcomes. Although there is good internal reliability for the twin difference measures for both outcome variables, the internal reliability of the parenting predictor is somewhat lower (alpha=.65). Perhaps observed parenting scores are affected by random factors unrelated to "true" parenting scores; for example, parent teaching behaviors may improve for the second twin tested due to greater familiarity with the task, artificially creating the impression that the second twin receives better parenting in general. Because error in predictor variables can bias outcomes, generally attenuating estimates downward (Bound et al., 2001), the low reliability on this measure is particularly problematic.

However, the robustness of the estimates to measurement errors can be tested by performing an errors-in-variables regression (*eivreg* in Stata). This procedure adjusts coefficient estimates to account for the fact that some variables are measured with error. Estimates using *eivreg* were substantively similar to the main results.[1] The coefficients on the parenting variable were far from significant even when estimates were adjusted for error in the independent measures, for both the cognitive and socioemotional outcomes (b=.005, n.s. and b=.065, n.s. respectively; results available upon request). This suggests that even if parenting and outcome variables were measured without error, the conclusions of the study would hold.

Moreover, there is some evidence that the results presented here are trustworthy. Several variables *do* reliably predict differences in child outcomes at 2 years, even in the twin differences

[1] The equations used for the EIV estimation were slightly different from the equations used in the rest of the analysis (e.g., some missing dummy indicators tolerated by normal regression procedures had to be dropped when *eivreg* was used due to near-perfect collinearity, and the constant could not be suppressed using *eivreg*.) However, I conducted a set of regressions using the same regressors as in *eivreg*, but with the same procedures used in the rest of the paper. These analyses suggest that cleansing measurement error does not change the study's conclusions.

models. For instance, male sex is negatively related to both cognitive and socioemotional development at 24 months (Tables 4 and 5). These results are consistent with past research (Baillargeon et al., 2007; Carson, Wagner, & Schultz, 1987; Galsworthy et al., 2000).

Furthermore, differences in twins' birth weights predict differences in their cognitive development at 24 months. Past research has linked low birth weight with cognitive decrements, particularly in early childhood (Breslau et al., 1996; Tong, Baghurst, & McMichael, 2006). Previous twin studies have also indicated that higher birth weights are associated with cognitive advantages as well (Asbury et al., 2006). Finally, differences in child age are also predictive of differences in cognitive development. Since the Bayley scales are sensitive to child age, this is consistent with expectations (Andreassen & Fletcher, 2007). The fact that the twin differences models are not insensitive to *all* predictors suggests that the null finding on parenting is real.

In addition, the diminished relationship between parenting and the various outcomes occurs not only because of the larger standard errors that might be expected as a result of using twin difference estimation. Rather, the *magnitude* of the coefficients drops substantially when the twin differences models are used. For instance, the coefficient on the parenting variable in the Bayley Mental Scale regression indicates that a one standard deviation improvement in parenting for one twin relative to the other is related to only a .004 standard deviation increase in cognitive performance at 24 months. While the standard error for this estimate is larger than the standard errors in the OLS models, a coefficient of this magnitude would not be significant even if it were judged against the smaller standard errors on the parenting coefficient found in the other models. Moreover, as noted earlier, the standard errors on both the main effects and the interaction terms are generally sufficiently precise that even effects considered "small" by Cohen's (1992) standards could be detected in the twin differences models.

Nonetheless, the discrepancy between these results and results found in earlier studies—including longitudinal studies using twin differences—begs the question of why these findings differ. Several explanations may be forwarded. This study uses observational ratings of parenting and socioemotional outcomes, and achievement tests for cognitive outcomes. Some previous studies that found parenting effects used parent-report measures for outcomes and/or for parenting variables (e.g., Asbury et al., 2006; Asbury et al., 2003). Using parent reports for both measures presents the threat of shared-source bias (Caspi et al., 2004).

By contrast, using parents to report their own parenting behaviors could produce results that are either more accurate or less accurate if a second source is used for the outcome measure. Parents have a better sense of their parenting behaviors on a daily basis than do NCATS coders, who view only a snapshot of parent behaviors. Thus, parent reports may give a more holistic view of their parenting practices. On the other hand, parents may not want to admit to differential treatment of their twins, and may therefore report less objectively than would a third party. Either condition may account for differences between this study and previous work, although they have different implications for which set of conclusions should be accepted.

Moreover, several studies that employed twin differences used bivariate correlations to test for relationships between parenting and subsequent outcomes (e.g., Asbury et al., 2006; Deater-Deckard et al., 2001). Other studies used multiple regression, but did not include any child-specific controls (e.g., Asbury et al., 2003). The present study controls for child-specific measures that may affect both parenting and development, like birth weight or child development measures at the baseline measurement. These results may therefore be more reliable.

A few prior studies, however, found statistically significant relationships between parenting and subsequent child outcomes despite employing baseline controls of child development and using third-party reporters of both parenting measures and child outcomes (e.g., Caspi et al., 2004). The magnitude of these results is also larger than those found in the current study. Absolute values of standardized regression coefficients on maternal expressed emotion measures in Caspi and colleagues' (2004) study, which linked parenting to antisocial behavior, range from .1 to .16. By contrast, the twin differences models in the present study found a standardized regression coefficient of .02 for the model linking parenting to socioemotional outcomes. While the measures used in these studies differ, the disparate results suggest that the present study requires replication before its results are accepted.

While there is some evidence that this study's results are valid, further work could be done to address the objections to its conclusions raised here. Structural equation models could be employed to strip random variation from measurements and establish the relationships between the latent constructs of parenting and child cognitive and socioemotional development. Future work should also use twin difference models on other data sets with twin samples to test the relationship between parenting and child cognitive and socioemotional outcomes, controlling for baseline measures of child development and other child-specific variables like birth weight. This program of research would help further knowledge about the associations between parenting in infancy and subsequent child cognitive and socioemotional development.

References

Allhusen, V., Belsky, J., Kersey, H. B., Booth, C., Bradley, R., Brownell, C. A., et al. (2003). Do children's attention processes mediate the link between family predictors and school readiness? [Article]. *Developmental Psychology, 39*(3), 581-593.

Allison, P. D. (1990). Change Scores as Dependent Variables in Regression Analysis (Vol. 20, pp. 93-114): American Sociological Association.

Andreassen, C., & Fletcher, P. (2005). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) methodology report for the nine-month data collection. Volume 1: Psychometric Characteristics*: National Center for Education Statistics.

Andreassen, C., & Fletcher, P. (2007). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B): Psychometric report for the 2-year data collection, methodology report*: National Center for Education Statistics.

Asbury, K., Dunn, J. F., & Nomin, R. (2006). Birthweight-discordance and differences in early parenting relate to monozygotic twin differences in behaviour problems and academic achievement at age 7. *Developmental Science, 9*(2), F22-F31.

Asbury, K., Dunn, J. F., Pike, A., & Plomin, R. (2003). Nonshared environmental influences on individual differences in early behavioral development: A monozygotic twin differences study. *Child Development, 74*(3), 933-943.

Baillargeon, R. H., Zoccolillo, M., Keenan, K., Cote, S., Perusse, D., Wu, H. X., et al. (2007). Gender differences in physical aggression: A prospective population-based survey of children before and after 2 years of age. [Article]. *Developmental Psychology, 43*(1), 13-26.

Barnard, K. E., Hammond, M. A., Booth, C., Bee, H. L., Mitchell, S. K., & Spieker, S. (1989). Measurement and meaning of parent-child interaction. In F. J. Morrison, C. Lord & D. P. Keating (Eds.), *Applied Developmental Psychology* (3rd ed., pp. 40-80). New York: Academic Press, Inc.

Bee, H. L., Barnard, K. E., Eyres, S. J., Gray, C. A., Hammond, M. A., Spietz, A. L., et al. (1982). Prediction of IQ and language skill from perinatal status, child performance, family characteristics, and mother-infant interaction. *Child Development, 53*(5), 1134-1156.

Bornstein, M. H. (1985). How Infant and Mother Jointly Contribute to Developing Cognitive Competence in the Child (Vol. 82, pp. 7470-7473): National Academy of Sciences.

Bornstein, M. H. (2002). Parenting infants. In M. H. Bornstein (Ed.), *Handbook of Parenting: Children and Parenting* (Vol. 1, pp. 3-43). Mahwah, NJ: Lawrence Erlbaum Associates.

Bound, J., Brown, C., & Mathiowetz, N. (2001). Measurement error in survey data. In J. J. Heckman & E. Leamer (Eds.), *Handbook of Econometrics* (Vol. 5, pp. 3705-3843): Elsevier Science B.V.

Bradley, R. H., Caldwell, B. M., & Rock, S. L. (1988). Home-Environment and School Performance - a 10-Year Follow-up and Examination of 3 Models of Environmental Action. *Child Development, 59*(4), 852-867.

Bradley, R. H., Corwyn, R. F., Burchinal, M., McAdoo, H. P., & Coll, C. G. (2001). The home environments of children in the United States part II: Relations with behavioral development through age thirteen. [Article]. *Child Development, 72*(6), 1868-1886.

Breslau, N., Chilcoat, H., DelDotto, J., Andreski, P., & Brown, G. (1996). Low birth weight and neurocognitive status at six years of age. *Biological Psychiatry, 40*(5), 389-397.

Burt, S. A., McGue, M., Iacono, W. G., & Krueger, R. F. (2006). Differential parent-child relationships and adolescent externalizing symptoms: Cross-lagged analyses within a monozygotic twin differences design. *Developmental Psychology, 42*(6), 1289-1298.
Carson, D. K., Wagner, B. S., & Schultz, N. W. (1987). Temperament and gender-Correlates of toddler social competence. *Journal of Genetic Psychology, 148*(3), 289-302.
Caspi, A., Moffitt, T. E., Morgan, J., Rutter, M., Taylor, A., Arseneault, L., et al. (2004). Maternal expressed emotion predicts children's antisocial behavior problems: Using monozygotic-twin differences to identify environmental effects on behavioral development. *Developmental Psychology, 40*(2), 149-161.
Chase-Lansdale, P. L., Wakschlag, L. S., & Brooks-Gunn, J. (1995). A Psychological Perspective on the Development of Caring in Children and Youth - the Role of the Family. *Journal of Adolescence, 18*(5), 515-556.
Cohen, J. (1992). A power primer. [Article]. *Psychological Bulletin, 112*(1), 155-159.
Collins, W. A., Maccoby, E. E., Steinberg, L., Hetherington, E. M., & Bornstein, M. H. (2000). Contemporary research on parenting - The case for nature and nurture. [Review]. *American Psychologist, 55*(2), 218-232.
Cronbach, L. J., & Furby, L. (1970). How should we measure "change"--or should we? *Psychological Bulletin, 74*(1), 68-80.
Deater-Deckard, K., Bates, J. E., Dodge, K. A., & Pettit, G. S. (1996). Physical discipline among African American and European American mothers: Links to children's externalizing behaviors. [Article]. *Developmental Psychology, 32*(6), 1065-1072.
Deater-Deckard, K., Pike, A., Petrill, S. A., Cutting, A. L., Hughes, C., & O'Connor, T. G. (2001). Nonshared environmental processes in social-emotional development: an observational study of identical twin differences in the preschool period. *Developmental Science, 4*(2), F1-F6.
Duncan, G. J., Claessens, A., Huston, A. C., Pagani, L. S., Engel, M., Sexton, H., et al. (2007). School readiness and later achievement. [Article]. *Developmental Psychology, 43*(6), 1428-1446.
Duncan, G. J., Yeung, W. J., Brooks-Gunn, J., & Smith, J. R. (1998). How much does childhood poverty affect the life chances of children? *American Sociological Review, 63*(3), 406-423.
Elardo, R., Bradley, R., & Caldwell, B. M. (1977). Longitudinal study of relation of infants home environments to language development at age 3. *Child Development, 48*(2), 595-603.
Findji, F. (1993). Attentional abilities and maternal scaffolding in the first year of life. [Article]. *International Journal of Psychology, 28*(5), 681-692.
Galsworthy, M. J., Dionne, G., Dale, P. S., & Plomin, R. (2000). Sex differences in early verbal and non-verbal cognitive development. *Developmental Science, 3*(2), 206-215.
Gleason, J. B., & Ely, R. (2002). Gender differences in language development. In A. McGillicuddy-DeLisi & R. DeLisi (Eds.), *Biology, Society, and Behavior: The Development of Sex Differences in Cognition* (Vol. 21, pp. 127-154). Westport, CT: Ablex Publishing.
Ispa, J. M., Fine, M. A., Halgunseth, L. C., Harper, S., Robinson, J., Boyce, L., et al. (2004). Maternal intrusiveness, maternal warmth, and mother-toddler relationship outcomes: Variations across low-income ethnic and acculturation groups. *Child Development, 75*(6), 1613-1631.

Kelly, J. F., & Barnard, K. E. (2000). Assessment of parent-child interaction: Implications for early intervention. In J. P. Shonkoff & S. J. Meisels (Eds.), *Handbook of Early Childhood Intervention* (2nd ed., pp. 258-289). Cambrdige: Cambridge University Press.

Kelly, J. F., Morisset, C. E., Barnard, K. E., Hammond, M. A., & Booth, C. L. (1996). The influence of early mother-child interaction on preschool cognitive/linguistic outcomes in a high-social-risk group. *Infant Mental Health Journal, 17*(4), 310-321.

Kochanska, G., Friesenborg, A. E., Lange, L. A., & Martel, M. M. (2004). Parents' personality and infants' temperament as contributors to their emerging relationship. [Article]. *Journal of Personality and Social Psychology, 86*(5), 744-759.

Landry, S. H., Smith, K. E., MillerLoncar, C. L., & Swank, P. R. (1997). Predicting cognitive-language and social growth curves from early maternal behaviors in children at varying degrees of biological risk. [Article]. *Developmental Psychology, 33*(6), 1040-1053.

Liang, H. A., & Eley, T. C. (2005). A monozygotic twin differences study of nonshared environmental influence on adolescent depressive symptoms. *Child Development, 76*(6), 1247-1260.

Lytton, H., & Gallagher, L. (2002). Parenting twins and the genetics of parenting. In M. H. Bornstein (Ed.), *Handbook of Parenting: Children and Parenting* (2nd ed., Vol. 1, pp. 227-253). Mahwah, NJ: Lawrence Erlbaum Associates.

Maccoby, E. E. (2003). The gender of child and parent as factors in family dynamics. In A. C. Crouter & A. Booth (Eds.), *Children's Influence on Family Dynamics: The neglected side of family relationships* (pp. 191-206). Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Mistry, R. S., Biesanz, J. C., Chien, N., Howes, C., & Benner, A. D. (2008). Socioeconomic status, parental investments, and the cognitive and behavioral outcomes of low-income children from immigrant and native households. *Early Childhood Research Quarterly, 23*(2), 193-212.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). New York: McGraw-Hill, Inc.

Pike, A., Iervolino, A. C., Eley, T. C., Price, T. S., & Plomin, R. (2006). Environmental risk and young children's cognitive and behavioral development. *International Journal of Behavioral Development, 30*(1), 55-66.

Reiss, D., Neiderhiser, J. M., Hetherington, E. M., & Plomin, R. (2000). *The Relationship Code: Deciphering genetic and social influences on adolescent development*. Cambridge, M.A.: Harvard University Press.

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child Development, 72*(3), 748-767.

Tong, S. L., Baghurst, P., & McMichael, A. (2006). Birthweight and cognitive development during childhood. [Article]. *Journal of Paediatrics and Child Health, 42*(3), 98-103.

Turkheimer, E., & Waldron, M. (2000). Nonshared environment: A theoretical, methodological, and quantitative review. *Psychological Bulletin, 126*(1), 78-108.

Wacharasin, C. (2001). *Predicting child cognitive development in low-income families.* University of Washington.

Wacharasin, C., Barnard, K. E., & Spieker, S. J. (2003). Factors affecting toddler cognitive development in low-income families - Implications for practitioners. *Infants and Young Children, 16*(2), 175-181.

Wakschlag, L. S., & Hans, S. L. (1999). Relation of maternal responsiveness during infancy to the development of behavior problems in high-risk youths. *Developmental Psychology, 35*(2), 569-579.

Wooldridge, J. M. (2003). *Introductory Econometrics: A Modern Approach*. Singapore: Sout-Western.

Table 1. Descriptive statistics

	Full sample: Mean (sd)[1]	Twin sample: Mean (sd)[2]	Range
Child outcome dependent measures-2 years			
BSF-R-Mental scale	127.09 (10.65)	123.17 (10.62)**	92.35-174.14
Behavior Rating Scale	39.85 (8.25)	37.49 (8.42)**	11-55
9-month parenting predictors			
NCAT parent	34.77 (4.48)	34.60 (4.50)	15-49
Controls that vary between twins-9 months			
BSF-R-Mental scale	76.77 (9.78)	71.91 (9.02)**	32.04-131.17
Infant/Toddler Symptom Checklist	8.09 (3.83)	8.12 (4.12)	0-21
Behavior Rating Scale	25.95 (4.57)	24.44 (4.63)**	7-35
BSF-R-Motor scale	55.96 (9.40)	51.87 (9.07)**	21.16-87.10
NCAT child	15.51 (2.71)	14.9 (2.88)**	4-23
Birth weight	7.32 (1.31)	5.32 (1.28)**	.50-12.00
%male	.5096	.4973	
Child age at 2nd wave (months)	24.39 (1.19)	24.33 (1.02)*	20.1-38.2
Family fixed characteristics			
%white	.5458	.6186*	
%black	.1328	.1375	
%Hispanic	.2501	.1869*	
%Asian	.026	.017*	
%other	.045	.0397	
% parents married-9 months	.6776	.7441**	
Family income-9 months ($10K)	5.09 (4.41)	5.99 (4.89)**	.25-20
Mother age at birth	27.44 (6.19)	29.38 (5.96)**	15-50
Maternal education at birth	12.96 (2.84)	13.52 (2.74)**	0-17

1)N's at time 1 vary between 8012 and 9832. At time 2, N's vary between 8826 and 9458.
2)N's at time 1 vary between 1224 and 1516. At time 2, N's vary between 1376 and 1470.
3)N's in correlations vary between 88 and 120 for MZ pairs and 401 and 640 for DZ twins.
* significantly different from full sample measures at $p<.05$; ** significantly different at $p<.01$

Table 2. Reliability coefficients

	Full sample[2]	Twin subsample	Twin differences
Child outcome dependent measures (2 years)			
Bayley Mental	.88	.89	.87
Interviewer Behavior Ratings	-	.91	.88
Parenting predictors (9 months)			
NCATS parent items	.68	.72	.65
Controls that vary between twins (9 months)			
Bayley Mental	.81	.71	.62
Interviewer Behavior Ratings	.79	.78	.75
Infant/Toddler Symptom Checklist	.63	.64	.58
Bayley Motor	.94	.87	.83
NCATS child items	.62	.64	.56

[2] From NCES; see Andreassen & Fletcher (2005, 2007)

Table 3. Correlations

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Mental 2	1		.48*	.00	.08*	-.15*	.04	.04	.17*	.06	.04	.01					
BRS 2	2	.53*		.01	.00	-.17*	.01	.04	.10*	.04	-.03	.00					
NCATS parenting	3	.18*	.10*		.00	-.02	.04	.04	-.06	-.03	.14*	-.06					
Birth weight	4	.14*	.08*	.06*		.10*	.01	.10*	.11*	.06	.11*	-.02					
Male	5	-.17*	-.15*	-.03*	.10*		-.04	.05	.03	.05	-.04	.02					
Child age-2 yr	6	.20*	.06*	.01	-.02	.01		.05	.03	-.07*	.02	-.01					
Mental 1	7	.20*	.11*	.17*	.13*	-.02*	.19*		.22*	.21*	-.06	-.01					
Motor 1	8	.14*	.08*	.09*	.12*	.01	.16*	.72*		.22*	.00	-.01					
BRS 1	9	.16*	.21*	.10*	.09*	-.04*	.02	.30*	.26*		.05	-.09*					
NCATS child	10	.09*	.04*	.24*	.03*	-.03*	.03*	.05*	.01	.03*		-.04					
ITSC 1	11	-.05*	-.06*	-.03*	-.02*	.03*	.02*	.02	.05*	-.08*	.00						
Income 1	12	.23*	.11*	.24*	.10*	.03*	-.05*	.02*	-.05*	.02*	.04*	-.11*					
Married 1	13	.15*	.10*	.19*	.10*	-.02	-.05*	.01	-.07*	.00	.01	-.07*	.38*				
Maternal age	14	.11*	.09*	.17*	.07*	-.01	-.05*	-.03*	-.10*	-.01	.01	-.09*	.43*	.42*			
Maternal education	15	.27*	.13*	.29*	.06*	-.02*	-.06*	.02*	-.03*	.06*	.05*	-.07*	.51*	.35*	.41*		
Black	16	-.11*	-.02*	-.07*	-.13*	.01	.00	-.02*	.07*	-.01	.01	.06*	-.19*	-.31*	-.14*	-.07*	
Hispanic	17	-.20*	-.06*	-.15*	.00	.00	.04*	-.01	-.02*	-.05*	.00	.01	-.22*	-.12*	-.12*	-.38*	-.23*

Correlations for the full sample are below the diagonal; correlations between twin differences are above the diagonal. Because variables 12-18 (Income-Hispanic) are shared between twins, I only report correlations for the first 11 variables for twin differences.

BRS=Behavior Rating Scale; NCATS=Nursing Child Assessment Teaching Scale; ITSC=Infant/Toddler Symptom Checklist

Table 4. Bayley Mental Scale-2 year

	Bivariate OLS Models		OLS with demographic controls	Fully-controlled OLS models		Twin differences
	(1)	(2)	(3)	(4)	(5)	(6)
	Full sample	Twin sample	Full sample	Full sample	Twin sample	Twin sample
NCATS parenting score-9 months	0.177**	0.219**	0.072**	0.034*	0.094**	0.004
	(0.015)	(0.033)	(0.015)	(0.016)	(0.035)	(0.047)
Birth weight-lbs			0.096**	0.078**	0.097**	0.118**
			(0.008)	(0.008)	(0.028)	(0.041)
Child male			-0.362**	-0.345**	-0.360**	-0.235**
			(0.026)	(0.026)	(0.055)	(0.053)
Child age-months			0.181**	0.162**	0.206**	0.046*
			(0.016)	(0.015)	(0.036)	(0.019)
Family income-10,000s			0.022**	0.022**	0.016*	
			(0.005)	(0.005)	(0.007)	
Married			0.047	0.053+	-0.014	
			(0.032)	(0.031)	(0.099)	
Maternal age-birth			-0.007*	-0.005+	-0.008	
			(0.003)	(0.003)	(0.006)	
Maternal education-birth			0.058**	0.057**	0.078**	
			(0.007)	(0.007)	(0.015)	
Black, non-Hispanic			-0.286**	-0.283**	-0.514**	
			(0.041)	(0.041)	(0.142)	
Hispanic			-0.308**	-0.309**	-0.179+	
			(0.041)	(0.042)	(0.103)	
Bayley Mental Scale-9 months				0.107**	0.110*	0.018
				(0.021)	(0.048)	(0.067)
Bayley Motor Scale-9 months				0.007	0.052	0.129*
				(0.020)	(0.053)	(0.064)
Behavior Rating Scale-9 months				0.074**	0.021	-0.031
				(0.015)	(0.032)	(0.031)
NCATS child score-9 months				0.049**	0.018	0.013
				(0.015)	(0.028)	(0.030)
ITSC-9 months				-0.026*	-0.033	0.009
				(0.011)	(0.031)	(0.033)
Missing variable dummies	X	X	X	X	X	X
Constant	0.018	0.021	-5.499**	-4.966**	-6.131**	
	(0.021)	(0.043)	(0.427)	(0.411)	(0.924)	
Observations	8944	1161	8944	8944	1161	535
R-squared	0.032	0.047	0.204	0.228	0.275	0.062

Jackknife standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 5. Behavior Rating Scale-2 years

	Bivariate OLS Models		OLS with demographic controls	Fully-controlled OLS models		Twin differences
	(1)	(2)	(3)	(4)	(5)	(6)
	Full sample	Twin sample	Full sample	Full sample	Twin sample	Twin sample
NCATS parenting score-9 months	0.101**	0.127**	0.053**	0.029+	0.048	0.020
	(0.015)	(0.035)	(0.014)	(0.016)	(0.044)	(0.042)
Birth weight-lbs			0.065**	0.048**	0.012	0.046
			(0.009)	(0.009)	(0.028)	(0.043)
Child male			-0.308**	-0.286**	-0.322**	-0.331**
			(0.023)	(0.023)	(0.055)	(0.059)
Child age-months			0.063**	0.056**	0.091*	-0.000
			(0.015)	(0.014)	(0.037)	(0.015)
Family income-10,000s			0.010**	0.010**	0.002	
			(0.004)	(0.004)	(0.008)	
Married			0.088*	0.099**	0.143	
			(0.036)	(0.035)	(0.116)	
Maternal age-birth			0.003	0.004	-0.008	
			(0.003)	(0.003)	(0.006)	
Maternal education-birth			0.022**	0.020**	0.056**	
			(0.006)	(0.006)	(0.014)	
Black, non-Hispanic			0.057	0.065	-0.137	
			(0.047)	(0.045)	(0.132)	
Hispanic			-0.004	0.004	-0.191+	
			(0.051)	(0.047)	(0.114)	
Bayley Mental Scale-9 months				0.031	0.079	0.040
				(0.024)	(0.051)	(0.066)
Bayley Motor Scale-9 months				0.007	0.063	0.101
				(0.023)	(0.054)	(0.071)
Behavior Rating Scale-9 months				0.178**	0.125**	-0.016
				(0.018)	(0.036)	(0.037)
NCATS child score-9 months				0.013	0.027	-0.032
				(0.018)	(0.038)	(0.037)
ITSC-9 months				-0.034*	-0.063+	0.001
				(0.013)	(0.035)	(0.027)
Missing variable dummies	X	X	X	X	X	X
Constant	0.023	0.009	-2.299**	-2.042**	-2.720**	
	(0.026)	(0.039)	(0.393)	(0.358)	(0.967)	
Observations	8944	1150	8944	8944	1150	529
R-squared	0.010	0.015	0.061	0.098	0.145	0.058

Jackknife standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 6. Bayley Mental Scale-2 year: Interactions-Black

	Fully-controlled OLS models		Twin differences
	(1) Full sample	(2) Twin sample	(3) Twin sample
NCATS parenting score-9 months	0.038*	0.112**	0.000
	(0.017)	(0.038)	(0.048)
NCATS parenting*Black	-0.034	-0.130+	0.008
	(0.034)	(0.071)	(0.107)
Black, non-Hispanic	-0.289**	-0.539**	
	(0.041)	(0.142)	
Other controls	X	X	X
Missing variable dummies	X	X	X
Observations	8944	1159	534
R-squared	0.229	0.276	0.068

Jackknife standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 7. Bayley Mental Scale-2 year: Interactions-Below 185% of poverty line

	Fully-controlled OLS models		Twin differences
	(1) Full sample	(2) Twin sample	(3) Twin sample
NCATS parenting score-9 months	0.057**	0.096*	0.015
	(0.019)	(0.042)	(0.065)
NCATS parenting*Below 185% of poverty line	-0.051*	-0.005	-0.024
	(0.022)	(0.058)	(0.078)
Below 185% of poverty line	-0.185**	-0.269**	
	(0.032)	(0.074)	
Other controls	X	X	X
Missing variable dummies	X	X	X
Constant	-4.863**	-5.726**	
	(0.413)	(0.906)	
Observations	8944	1161	535
R-squared	0.229	0.280	0.062

Jackknife standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 8. Behavior Rating Scale-2 years: Interactions-Below 185% of poverty line

	Fully-controlled OLS models		Twin differences
	(1) Full sample	(2) Twin sample	(3) Twin sample
NCATS parenting score-9 months	0.028+	0.056	0.129*
	(0.016)	(0.044)	(0.050)
NCATS parenting*Below 185% of poverty line	0.001	-0.022	-0.241**
	(0.025)	(0.073)	(0.086)
Below 185% of poverty line	-0.078*	-0.081	
	(0.033)	(0.081)	
Other controls	X	X	X
Missing variable dummies	X	X	X
Constant	-2.007**	-2.570**	
	(0.363)	(0.961)	
Observations	8944	1150	529
R-squared	0.098	0.145	0.077

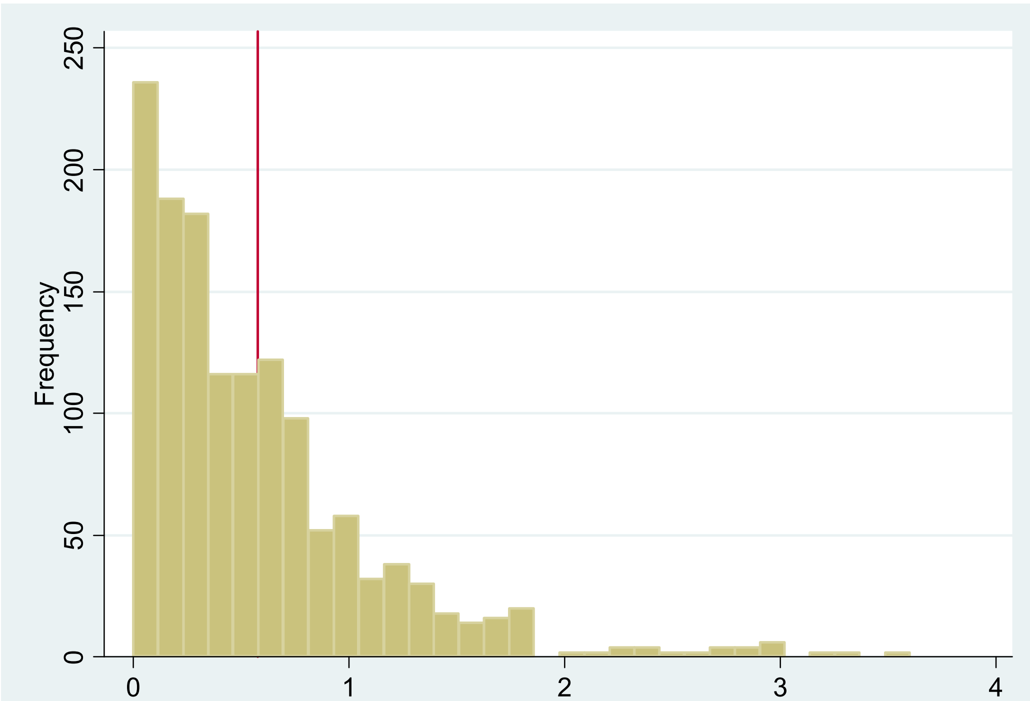
Jackknife standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%

Table 9. Behavior Rating Scale-2 years: Mean split temperament interactions

	Fully-controlled OLS models		Twin differences
	(1) Full sample	(2) Twin sample	(3) Twin sample
NCATS parenting score-9 months	0.031	0.074+	0.011
	(0.022)	(0.044)	(0.039)
NCATS parenting*below-average self-regulation	-0.003	-0.056	-0.155**
	(0.031)	(0.052)	(0.045)
Below-average self-regulation	-0.059*	-0.150*	-0.012
	(0.029)	(0.074)	(0.055)
Constant	-2.017**	-2.587**	
	(0.354)	(0.965)	
Observations	8944	1148	528
R-squared	0.098	0.146	0.069

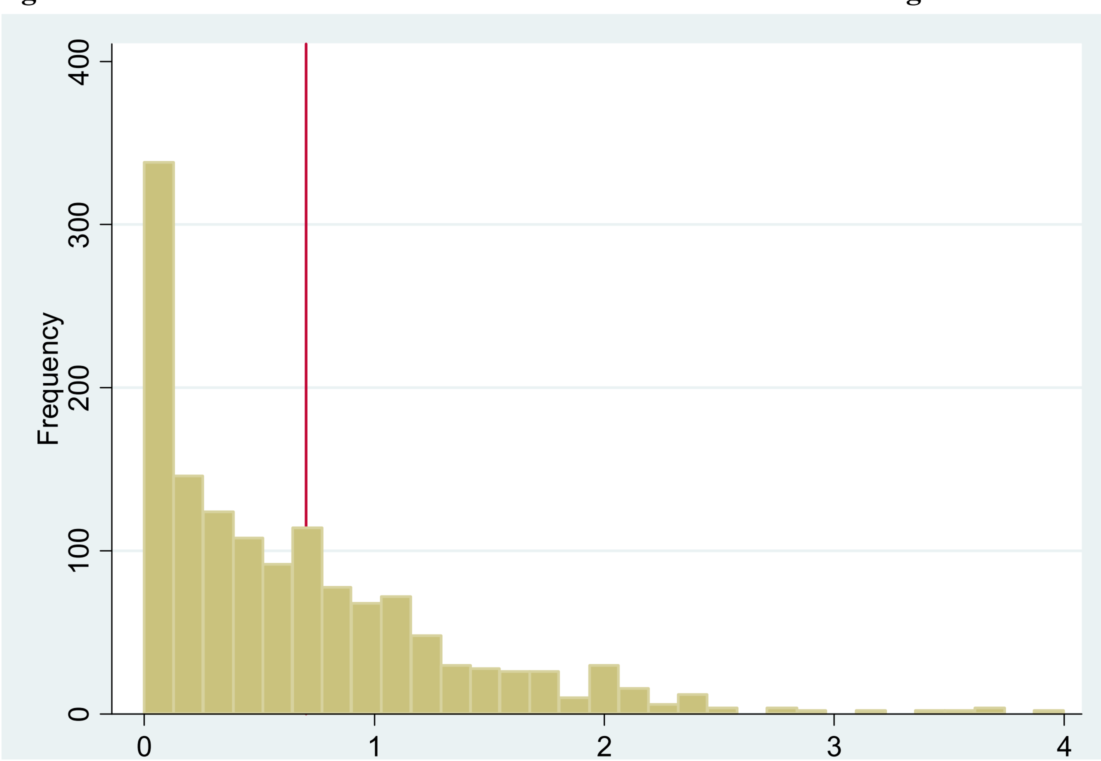
Jackknife standard errors in parentheses
+ significant at 10%; * significant at 5%; ** significant at 1%

Figure 1. Distribution of absolute twin differences. Bayley Mental Scale scores-2 years.



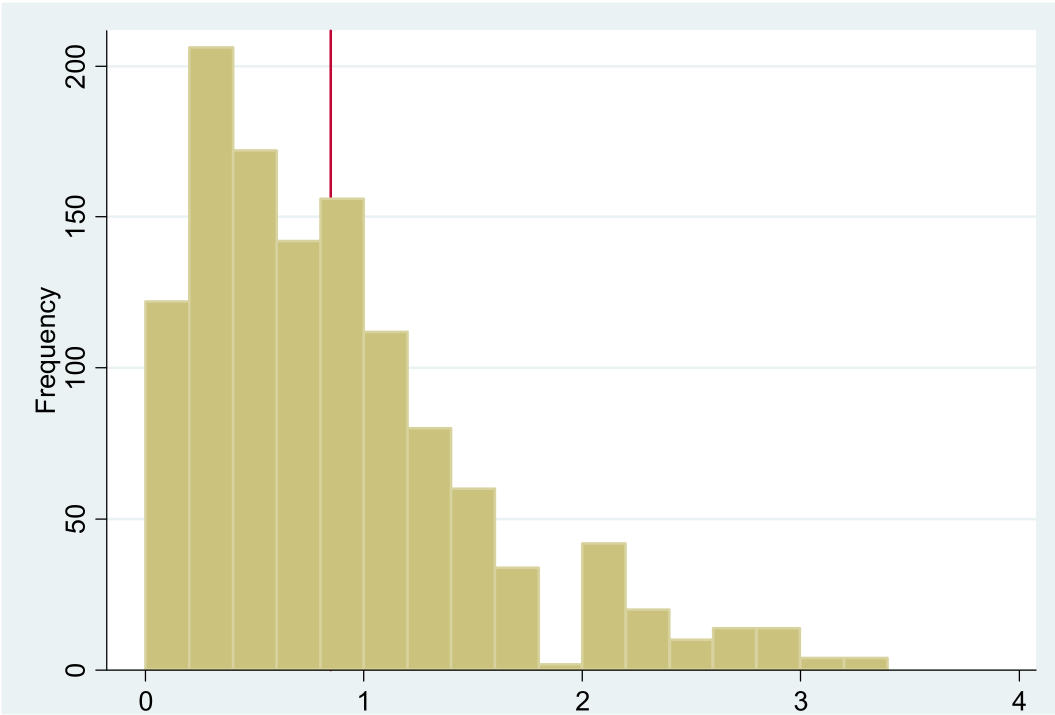
Red line represents mean absolute difference between twins=.58 full sample standard deviations.

Figure 2. Distribution of absolute twin differences. Behavior Rating Scale scores-2 years.



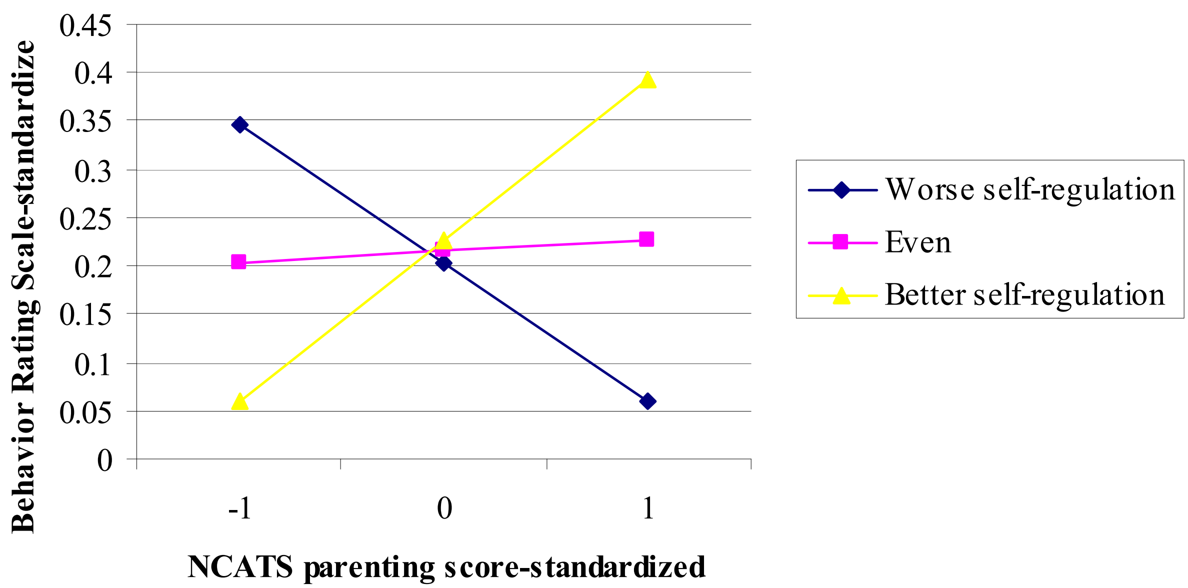
Red line represents mean absolute difference between twins=.70 full sample standard deviations.

Figure 3. Distribution of absolute twin differences. NCATS parenting scores-9 months.



Red line represents mean absolute difference between twins=.85 full sample standard deviations.

Figure 4. Predicted Behavior Rating Scale scores-2 years. Interaction between NCATS parenting score (standardized) and self-regulation problems (mean split) using twin differences.



Predicted values are for females, of sample-mean age (24.33 months) and birth weight (5.32 lbs), with scores of 0 on the standardized lagged child development measures.

Worse self-regulation: focus child has below-average self-regulation skills; the twin has above-average self-regulation skills.

Better self-regulation: focus child has above-average self-regulation skills; the twin has below-average self-regulation skills.

Even cases: both twins have above-average self-regulation skills, or both twins have below-average self-regulation skills.