



Published in final edited form as:

Dev Sci. 2011 January ; 14(1): 112–124. doi:10.1111/j.1467-7687.2010.00969.x.

A Longitudinal Analysis of Anger and Inhibitory Control in Twins from 12–36 Months of Age

Jeffrey R. Gagne and

Department of Psychology, University of Wisconsin, 1202 West Johnson Street, Madison, WI 53706-1611; phone: (608) 263-4735; fax: (608) 262-4029

H. Hill Goldsmith

Department of Psychology, University of Wisconsin, 1202 West Johnson Street, Madison, WI 53706-1611; phone: (608) 263-4735; fax: (608) 262-4029

Jeffrey R. Gagne: jgagne@wisc.edu; H. Hill Goldsmith: hhgoldsm@wisc.edu

Abstract

Inhibitory control (IC) is a dimension of child temperament that involves the self-regulation of behavioral responses under some form of instruction or expectation. Although IC is posited to appear in toddlerhood, the voluntary control of emotions such as anger begins earlier. Little research has analyzed relations between emotional development in infancy and later emerging IC. We examined phenotypic associations and genetic and environmental influences on parent- and laboratory-assessed anger and IC in a twin sample from 12 to 36 months of age. Typically, twins with low levels of IC had high levels of anger. Behavioral genetic findings confirmed significant genetic influences on anger and IC as assessed by parents, and on lab-based anger assessments. Shared environmental factors contributed to twin similarity on lab-assessed anger and IC at 36 months. Phenotypic covariance between anger and IC was largely due to overlapping genetic factors for parent ratings, and environmental factors in the laboratory.

Inhibitory control (IC) is a feature of child temperament involving the self-regulation of behavioral responses under some form of instruction or expectation (Goldsmith, 1996; Kochanska, Murray, Jacques, Koenig, & Vandegest, 1996; Rothbart, 1989a; Rothbart, Ahadi, Hershey & Fisher, 2001). Classic examples of IC include the ability to refrain from touching a hot stove, or from reaching for a cookie jar on a shelf, in the presence of an adult who has voiced a warning. Children who develop typical levels of IC are able to successfully inhibit behavior when necessary. Conversely, low IC often results in impulsive behavior. Early IC is significant because children with typically developing IC have fewer cognitive difficulties, more stable socio-emotional development, and less behavioral maladjustment.

Inhibitory Control and Early Temperament

According to many theorists, temperament describes early appearing behavioral and emotional tendencies, refers to individual differences in both reactivity and self-regulation, is considered relatively stable within broad developmental periods, tends to form the emotional core of personality traits, and is generally considered to be strongly biologically influenced (Goldsmith et al., 1987; Reed, Pien & Rothbart, 1984; Rothbart & Ahadi, 1994; Rothbart & Derryberry, 1981). IC abilities become apparent around two years of age, and

continue to develop in early childhood (Kochanska et al., 1996; Rothbart, 1989a). Many temperament researchers consider IC to be a major component of effortful control (EC), a less differentiated, earlier appearing temperament factor that involves the child's ability to engage in self-regulation (Eisenberg et al., 2001; 2004; Lemery-Chalfant, Doelger, & Goldsmith, 2008). EC is theorized to appear between 6 and 12 months of age (Kochanska, Murray, & Harlan, 2000; Rothbart & Ahadi, 1994), and is involved in the voluntary control of attentional processes and emotions (Derryberry & Rothbart, 1997; Rothbart, 1989a, 1989b; Rothbart & Ahadi, 1994; Rothbart & Bates, 1998).

The exercise of IC in toddlerhood requires children to actively regulate their behavior and inhibit acts that might otherwise lead to pleasurable outcomes (Kochanska, Murray, & Coy, 1997). Because IC overlaps with processes related to emotional control, it is plausible that patterns of early emerging emotional expression could be related to this later IC development. Recent evidence supports an indirect link between inhibition to novelty behavior (fear) in infancy, and IC in the toddler and early preschool years (Aksan & Kochanska, 2004). Other than for fear, longitudinal, developmental analyses of early affective dimensions related to IC are rare. Anger is another affective dimension that may influence IC. Do children who exhibit appropriate levels of anger in infancy also exhibit typical patterns of IC in early childhood? Perhaps angrier children have less behavioral control. Examining the association between anger and IC in infancy and early childhood can address this question and elucidate the developmental course of IC and the broader self-regulation of behavior.

Longitudinal Approaches

Longitudinal investigations have examined continuity and change in IC. When voluntary motor activation begins to develop in infancy, children approach exciting stimuli with little active inhibition. As the first year of life continues, infants show increased response latencies to approach an attractive toy (Reed et al., 1984). When IC emerges in the second year, children begin employing it more frequently as a behavioral strategy. IC behavior is thought to increase concurrently with the development of attention and verbal ability during the toddler and preschool years (Kochanska et al., 1997; Rothbart, 1989b). Scores on parent and observer ratings of IC generally improve across late infancy and early childhood (Kochanska et al., 2000). Success on lab-based IC tasks increases throughout childhood and then tapers off in adulthood (Williams, Ponesse, Schachar, Logan, & Tannock, 1999) although rank order within the group remains fairly stable. Thus, despite substantial mean level change, considerable stability of individual differences holds.

The complexity of IC behaviors also predictably changes across development. When IC begins to emerge, it is exemplified by the ability to alter behavior under instruction. In early childhood, IC is most often indexed by delay ability, the ability to delay a dominant response (Carlson, Davis, & Leach, 2005). When young children are asked to wait, they often become stimulus bound and have difficulties diverting their behavior (Kopp, 1982). Those children who have developed appropriate levels of IC are able to sustain waiting periods, whereas those with lower levels whine or engage in the behavior they are being encouraged to inhibit. As children mature, IC behaviors become more cognitively complex and are characterized by the ability to differentiate between multiple competing stimuli. The capacity to engage in response inhibition is the prototypical index of IC in older children. This capacity is typified by the suppression of a dominant response and engagement in an alternate response. Emotional expression may be more salient in the early development of IC than later, when complex executive functioning and attention skills are more involved in the exhibition of IC behavior. This is particularly apparent in the ways that IC is assessed in older children.

Contrary to IC, anger expression is present in infants as young as 7 months of age and laboratory-induced anger responses are considered normative in infancy (Stenberg, Campos, & Emde, 1983). In the context of increasing behavioral control, anger responses in childhood may be more atypical and indicative of qualitative differences between infancy and post-infancy anger expression. It is unclear if infancy and preschool anger will be similarly predictive of IC in early childhood. If the expression of anger changes across early development from a normative to a more atypical behavioral pattern, it is possible that infant anger may predict normative levels of IC, whereas preschool anger may predict lower levels of IC. Analyses of longitudinal samples that extend from infancy through toddlerhood can address these issues. Longitudinal analyses also allow for the elucidation of stability and change in early anger and IC.

The Etiology of Anger and Inhibitory Control

It is important to understand why some young children are able to control and inhibit their emotions and behavior better than others. Behavioral genetic methods address this issue of etiology. Previous twin research indicates that IC at 2 years of age (Gagne & Saudino, in press) and anger/frustration (Deater-Deckard, Petrill, & Thompson, 2007; Goldsmith, Buss, & Lemery, 1997) in childhood are genetically influenced. Related investigations in children and adults find genetic variance is present in EC (Goldsmith et al., 1997; Lemery-Chalfant et al., 2008; Yamagata, Takahashi, Kijima, Maekawa, Ono, & Ando, 2005). A behavioral genetic investigation of inhibition in 5–6 year-old children (Groot, de Sonnevile, Stins, & Boomsma, 2004) suggests that individual differences in IC might be explained by genetic factors, but model-fitting analyses could not distinguish between genetic versus shared environmental effects. Given these findings, it is reasonable to predict that individual differences in anger during infancy and early childhood, as well as IC in early childhood, will be associated with genetic variance. Even if these associations exist and the traits are heritable, the question of whether genetic factors contribute to the covariance between anger and IC still remains unanswered. Multivariate genetic analyses can clarify the factors that contribute to individual differences in anger and IC across age. That is, multivariate genetic analyses can address each of the three questions (are anger and IC phenotypically associated; are anger and IC heritable; and does the phenotypic association have a genetic basis?) simultaneously and in longitudinal perspective.

Most twin studies of temperament rely on parent reports for behavioral assessment; however, using parent ratings as the sole basis for inferences on the genetics of child temperament can introduce the potential for bias. Heritability estimates could be biased upward if parents contrast (inappropriately magnify) the respective behavioral styles of their dizygotic (DZ) twins, but not their monozygotic (MZ) twins (Saudino, 2003a; Saudino, Cherny, & Plomin, 2000). Specifically, twin studies that employ parent ratings of temperament often evince patterns of “too low” DZ twin correlations, suggesting that DZ twins are less similar than a randomly paired dyad (Neale & Stevenson, 1989; Plomin et al., 1993; Saudino, 2003a). DZ twins share on average 50% of their genetic material; therefore, if temperament is influenced by additive genetic factors, DZ twin correlations should not be less than half the magnitude of MZ twin correlations. The typically high correlations between MZ twins coupled with these very low (sometimes even zero or negative) correlations for DZ twins results in inflated estimates of genetic variance unless the genetic estimates are restricted by parameters of a model.

A special issue in *Infant Behavior and Development* (2003) focused on this topic, and most contributors agreed that these contrast effects and the validity of parent reports were an important consideration in temperament research and that some parent assessments of temperament were more susceptible to these biases than others (Goldsmith & Hewitt, 2003; Hwang & Rothbart, 2003; Saudino, 2003a; Saudino, 2003b; Seifer, 2003). Rating systems

that do not rely on global judgments of behavior and focus on specific, concrete behaviors show almost no evidence of contrast effects. Unfortunately, most past twin research on temperament has employed global parent assessments of behavior. In addition, a fairly extensive literature documents the lack of strong agreement between parent and laboratory ratings of temperament (Mangelsdorf, Schoppe, & Buur, 2000). This pattern of low convergence between parent and observer ratings may result from parental bias, but it also may reflect the reliability and validity of the more objective measures (Goldsmith & Hewitt, 2003). Therefore, it is important to use multiple sources of information about participants' behavior in a quantitative genetic analysis to more comprehensively investigate the etiology of the behavior under study (Saudino, 2005).

Employing both parent and observer ratings may provide a more accurate estimation of genetic and environmental factors that influence anger and IC than relying on parent ratings alone. Examining both parent and observer assessments will also allow for the evaluation of covariation between the two modalities and clarify whether contrast effects are operating on these traits. The finding of genetic variance for anger and IC would contribute toward understanding their etiology and toward understanding sources of early self-regulation.

The Present Study

This research examines phenotypic associations and genetic and environmental influences on parent- and laboratory-assessed anger and IC longitudinally, from 12 to 36 months of age. Multivariate behavioral genetic model-fitting analyses determined the relative contributions of genetic and environmental influences to individual differences in laboratory and parent ratings of anger and IC. These analyses also evaluated genetic and environmental contributions to the covariance between and across the different assessments of anger and IC. The comparison of nested models systematically tests hypotheses regarding genetic and environmental influences on each trait, continuity of these influences over time, and genetic and environmental covariance between traits.

Method

Participants

Participants included 735 (261 monozygotic or MZ twins, 474 dizygotic or DZ twins) children assessed for anger and IC by parents at 12 and 36 months, 846 (289 MZ, 557 DZ) children assessed for anger in the laboratory at 12 months, 1000 (331 MZ, 669 DZ) children assessed for anger in the laboratory at 36 months, and 1021 (346 MZ, 675 DZ) children assessed for IC in the laboratory at 36 months. These sample sizes refer to number of individuals rather than twin pairs. The sample was selected from families living in the greater Madison, Wisconsin, area who participated in a longitudinal twin study of social and emotional development from infancy to early childhood. Although this was a longitudinal investigation, additional families were recruited cross-sectionally at each study time point to boost sample size. Therefore, there are different sample sizes listed for laboratory assessments that occurred at different ages. In addition, fewer families completed parent reports, resulting in a smaller sample size for these questionnaire data. Although many longitudinal twin analyses utilize only participants who have complete data for each age under study, recent statistical practice has recognized the advantages of using all available data. The raw score-based data analyses employed in this study allow us to use all available data from each assessment occasion.

Twin pairs were recruited through a variety of methods including Wisconsin state birth records, mothers of twins clubs, television coverage, birth announcements in newspapers, doctors' offices, the Internet, and referrals from current participants. The children in the

study were born between the fall of 1991 and January of 2004. Twin zygosity was assessed with the Zygosity Questionnaire for Young Twins (Goldsmith, 1991). This method yields greater than 95% agreement with bloodtyping (Forget-Dubois et al., 2003; Price et al., 2000) and is practical and inexpensive. For 38 pairs, staff members were uncertain about a twin pair's zygosity after review of zygosity questionnaires, photographs and/or laboratory observations. In these cases, genotyping of a standard set of highly polymorphic markers verified the diagnosis.

Demographic data were based on the full sample. The sample was 50% female and approximately one-third of the children were categorized as MZ twins (35.0% MZ, 34.7% same-sex DZ, 30.3% opposite-sex DZ). The few twin pairs with no zygosity classification were not included in the quantitative genetic analyses. The racial composition of the families was 93.2% Caucasian, 3.4% African-American, 1.8% Asian-American, and 0.8% Native American (0.8% missing); 3.2% of the participants were classified as Hispanic. State birth records show that the percentage of Caucasian births in Dane county (the core recruitment area for the project) decreased from 85% in 1995 (these are the first public data available) to 76.6% in 2004, averaging about a one percent decrease per year. If we assume that this trend also held in 1991–1994, the average percentage of Caucasian births in Dane county was around 85% for the years sampled. Additionally, recruitment from surrounding counties in the catchment area most likely increased the proportion of Caucasian twins in our sample to 93.2% because those nearby counties typically had 98–99% Caucasian births during the same period. The average socioeconomic status of the twins was predominantly middle class according to the Hollingshead index (mean SES=46.7, *SD*=11.7) although there was considerable range (11–66) in the sample.

Procedure

The procedure involved two laboratory visits that occurred when the twins were 12 and 36 months of age. During the first visit, infant temperament was assessed using the Laboratory Temperament Assessment Battery Locomotor Version (Lab-TAB; Goldsmith & Rothbart, 1999), a comprehensive laboratory-based temperament assessment that includes behavioral episodes corresponding to specific dimensions of temperament. At 36 months, twins participated in a similar visit and the Preschool Version of the Lab-TAB assessed temperament (Goldsmith, Reilly, Lemery, Longley, & Prescott, 1995). Parents completed questionnaires about each child's temperament and family demographics and returned them by mail. Additional questionnaires were mailed to the parents when the twins were 22 months of age. The lab visits typically lasted less than two hours. During administration of the Lab-TAB episodes, children's behavior was videotaped and later rated by Lab-TAB coders. For all of the behavioral scoring, care was taken to ensure that reliability between coders and master coders (highly trained staff members) was maintained. Kappa values for all scored behaviors were required to be equal to or greater than .70. At least 10% of the cases were double coded by a master coder.

Laboratory Measures of Temperament

Infant anger was assessed at 12 months of age using the "Gentle Arm Restraint" (AR) and "Restraint in Car Seat" (CS) episodes from the Locomotor Lab-TAB. During the AR episode, the child was presented with a novel and interesting toy but prevented from playing with it by the parent. After the experimenter brought the toy into view and demonstrated how to play with it, parents were instructed to place their hands on the child's forearms from behind, and hold the child's arms gently to their sides for 30 seconds. Two separate trials were coded for this episode. In CS, the parent buckled the child into a standard car seat and then moved out of the child's view. The parent was instructed to leave the child in the seat for 30 seconds and refrain from comforting or speaking to the child. AR was again used as an

anger measure at 36 months. The Preschool Lab-TAB “End of the Line” (EL) episode was also used at this age (“Restraint in Car Seat” was no longer age-appropriate). During EL, the parent demonstrated an attractive toy and encouraged the child to play with it. After approximately one minute of play, the parent inexplicably took the toy away for 30 seconds.

Lab-based IC was assessed at 36 months using the “Dinky Toys” (DT) and “Snack Delay” (SD) episodes from the Preschool Lab-TAB. In the DT episode, the child was asked, on two trials, to select one toy out of a clear plastic container filled with several attractive trinkets, thereby inhibiting the urge to pick more than one toy or hoard several of them. During SD, the child was offered a snack (m&m candies or goldfish crackers) but was required to wait for a signal before eating it. The experimenter put the snack under a clear plastic cup and then rang a bell when it was permissible for the child to pick up the cup and eat the snack. There was one practice trial with no waiting time, and six test trials with different pause lengths (5s, 10s, 0s, 20s, 0s, 30s) before the experimenter rang the bell.

For the Lab-TAB episodes, item level data were converted to z-scores, and means, peaks, and latencies of behaviors were computed across trials if necessary. To approximate a more normal distribution, latency values were transformed to speed values by taking the inverse of the square root of all latency scores. This transformation allows latency scores to be weighted in the same direction as the other Lab-TAB variables.

Summary scores were formed for the 12-month AR and CS episodes using principal component analysis (PCA) of the two anger variables of facial anger and distress vocalizations. For 12-month AR, the first principal component accounted for 66% of the total variance and component loadings ranged from .64–.91. The first principal component for 12-month CS accounted for 81% of the variance and loadings ranged from .85–.95. Only one component was indicated by scree test for each episode. The AR and CS components were significantly correlated, and an overall infant anger composite was formed by computing the mean of the two episode scores.

A similar approach was followed for 36-month AR and EL, with anger variables of pulling free, kicking, arching back, intensity of struggle, facial anger and distress vocalizations used for the AR composite, and facial anger, bodily anger, and protest variables used for EL. To avoid contamination of anger measures by sadness/distress reactions in the AR, CS, and EL episodes, we separately scored a set of sadness/distress variables in these episodes. These sadness/distress measures were not used in the present analyses. Component loadings for AR at 36 months ranged from .59–.92 and the first principal component accounted for 54% of the variance. For EL, the first component accounted for 56% of the variance and loadings ranged from .63–.86. The mean of the AR and EL components formed the 36-month overall anger composite.

A summary score was formed for the DT episode using a PCA of the following variables: the child’s initial approach to the stimuli, latency to choose a toy, style of touching, following directions, and a global rating of impulsivity (the first principal component accounted for 51% of the variance). Component loadings for the first principal component of DT ranged from .44–.82

Forming the SD composite was not as straightforward as forming the composites just described. Rather than a single principal component, the initial PCA analyses for SD indicated three components (cumulatively accounting for 80% of the variance). The latency to eat and touch the snack, eating the snack early, and distract variables loaded on the first two components (all loadings greater than |.52|). The distract variables loaded negatively on the first component and positively on the second. This pattern suggested a rationale for forming a composite with the variables loading on the first and second components.

However, SD child prompt variables loaded on the third component (loadings ranged from .76–.92), which contributed to 18.5 % of the variance. The child prompt variables did not covary with the other variables, because a child prompting the experimenter to “ring the bell” so that the snack can be eaten preempts the other IC behaviors from occurring. That is, prompters could not engage in distraction from the experimenter or the stimuli, nor were they eating or touching the snack early. Nevertheless, child prompting is a rational candidate for indicating lower IC. In psychometric terminology, we believe that child prompting is “formative” rather than “reflective” in relation to the other variables that define the latent construct of inhibitory control in this episode. Therefore, we formed the SD summary score using an algorithm of the means of the latency to eat and touch the snack, eating the snack early, and distract variables (all of which were significantly intercorrelated) and the child prompt component from the PCA. The resulting SD summary scores were significantly correlated with the DT summary score, and an overall composite of observed IC was computed from the mean of these scores.

Parent Ratings of Temperament

The parent-rating questionnaires of temperament completed by the mothers were used in the present investigation. Infant temperament at 12 months of age was assessed by mothers using the Infant Behavior Questionnaire (IBQ; Rothbart, 1981). The IBQ asks parents to make judgments about their infant’s behavior during the previous week (Gartstein & Rothbart, 2003). The distress to limitation (anger) dimension¹ was included in the analyses. At age 36 months, mothers completed the Children’s Behavior Questionnaire (CBQ; Rothbart et al., 2001) to evaluate anger and IC. The CBQ calls for parents to judge behavior in certain contexts and is appropriate for children from 3 to 7 years of age (Rothbart et al., 2001). IBQ distress to limitation items focus on fussing and crying for infants and CBQ anger items highlight anger, irritation, and temper. The CBQ IC items focus on the inhibition of prepotent responses such as “sitting still when told,” “lowering one’s voice,” and “stopping activity when asked.” Prior to analyses with the parent-rated data, we used imputation to avoid biases due to missing data (Graham, 2009). SPSS Missing Value Analysis expectation maximization (EM) algorithms (Dempster, Laird, & Rubin, 1977) were used to estimate missing questionnaire data. Estimates of internal consistencies for the questionnaires ranged from .75 for CBQ anger to .83 for IBQ distress to limitation.

Descriptive Statistics and Phenotypic Correlations

Descriptive statistics and tests of mean sex differences were calculated for anger and IC variables. Phenotypic correlational analyses examined associations between the two temperament dimensions across age and assessment modality. Tests of mean differences and phenotypic correlations were corrected for the nested nature of twin data. Dyad-level correlations were computed using Griffin and Gonzalez’ GGEXCH programs for dyad level-data wherein the partners are exchangeable, such as twin pairs (Griffin & Gonzalez, 1995; O’Connor, 2004). Generalized Estimating Equation (GEE) models for correlated data tested for mean level sex differences (Liang & Zeger, 1986; Zeger & Liang, 1986).

¹There was some concern that the IBQ distress to limitations (anger) subscale was not entirely independent of the IBQ distress to novelty (i.e., fear) subscale at 12 months of age. These two IBQ subscales are correlated ($r = .30$) in the present study, and there is typically overlap between measures of negative affect in early childhood across both parent and laboratory ratings. We addressed the existence of a broad negative affect factor in infancy at the phenotypic level, using partial correlations. These partial correlations allowed us to determine whether the shared variance between IBQ distress to novelty and IBQ distress to limitations affected our conclusions about the IBQ distress to limitations variable’s association with IC. Partial correlations (controlling for IBQ distress to novelty) of IBQ distress to limitations with IC were comparable to the ordinary correlations between IBQ distress to limitations and IC that were reported in the manuscript. That is, there were no significant differences in the partial and full correlations between IBQ distress to limitations and the other variables, and differences in magnitude were negligible. Thus, we conclude that this potential bias or confound did not affect our results.

Twin Correlations

Intraclass correlations were computed for MZ and DZ twin pairs using a double entry procedure to provide an index of twin similarity. If MZ intraclass twin correlations exceed DZ values, genetic factors are implicated as contributing to individual differences on that trait. The foundation of multivariate approaches to behavioral genetics is the cross-twin, cross-trait correlation whereby twin A's score on one variable (e.g., anger) is correlated with twin B's score on the other variable (e.g., IC), and vice versa. Similar to intraclass correlations, cross-twin cross-trait correlations for MZ twins that exceed those of DZ twins imply that a proportion of the covariance between each trait is due to overlapping genetic influences. Twin covariances can be inflated by the variance due to sex; therefore, scores for all variables in the behavior genetic analyses were residualized for sex effects (McGue & Bouchard, 1984).

Multivariate Model-fitting Analyses

To further examine genetic and environmental influences on parent-assessed anger and IC, multivariate Cholesky decomposition models were applied to our twin data using Mx maximum-likelihood model-fitting procedures (Neale, 2003). To increase power, raw data were used in the Cholesky model-fitting analyses for anger and IC. Although univariate models can be used to estimate genetic and environmental variance singly, multivariate model-fitting allows for a more powerful test and partitions phenotypic *covariance* into genetic, shared, and nonshared environmental components (Schmitz, Cherny, & Fulker, 1998). A trivariate Cholesky decomposition model with laboratory-assessed anger at 12 and 36 months and lab-assessed IC at 36 months is depicted as a path diagram in Figure 1 (variables were entered in the model in this order). In this model, the observed phenotypic variance in the three traits is represented by the rectangles. The circles signify latent genetic and environmental variables. The latent variables A1, C1, and E1 represent the genetic, shared environmental, and nonshared environmental factors common to all three variables in the model; A2, C2, and E2 represent the genetic, shared environmental, and nonshared environmental factors common to the last two variables independent of the first; and A3, C3, and E3 are factors unique to the third variable. A similar Cholesky model was used for parent-assessed temperament.

Cholesky models are employed to estimate heritability, and shared and nonshared environmental variance for each individual phenotype; genetic and environmental correlations (i.e., r_g , r_c , r_e) between multiple variables; and the genetic and environmental contributions to phenotypic correlations between phenotypes. Genetic and environmental correlations indicate the degree to which genetic or environmental factors for one trait correlate with those on another trait. For example, the overlap of two moderately heritable traits could conceivably arise solely from environmental sources of variation. The extent to which overlapping genetic or environmental influences account for the phenotypic correlation between two traits can also be determined with multivariate analyses. Alternate models were tested and compared with the full model. Specifically, the A and C variances for each variable, and A, C, and E covariances between variables were dropped to determine if genetic and/or environmental variances and covariances were significant. Reduced models were assessed with the χ^2 difference test to determine if they were significantly different from the full model. A significant difference in χ^2 indicates a poorer fit, and that the parameter dropped from the model was significant and therefore must be retained.

Results

Gender Differences and Mean Change with Age

Table 1 lists the means and standard deviations of observer- and parent-rated anger and IC for both males and females. In general, males had higher levels of anger and lower levels of IC than females although these differences were typically significant for IC only. The preschool Lab-TAB was the only anger assessment that showed a significant difference between boys and girls ($z = 5.54, p < .0001$). Males were significantly lower in IC than females on the CBQ ($z = -5.39, p < .0001$), and this sex difference was replicated on the preschool Lab-TAB IC measure ($z = -7.93, p < .0001$). The effect sizes showed that males and females differed by approximately 38% of a standard deviation on preschool Lab-TAB anger, 41% on CBQ IC, and 55% on preschool Lab-TAB IC. It is important to highlight that lab-based anger at 12 months followed a normal distribution whereas, at 36 months of age, most of the children did not become angry during the anger episodes. This pattern indicates that anger assessed in the Lab-TAB may be a normative affective reaction in infancy, but non-normative in early childhood.

Phenotypic Correlations

Parents' ratings of anger were positively associated across 12 and 36 months of age (Table 2). However, the correlation between lab-assessed anger at 12 and 36 months was near zero. One possible explanation for this lack of association is that the Lab-TAB anger assessments become increasingly complex from 12 to 36 months. The 36-month measures include several new variables (e.g., pulling free, kicking, arching back, protest and restraint). To investigate this possibility, we formed a new set of 12- and 36-month anger composites in precisely the same way, using only facial anger and distress vocalizations variables. Even with these revised composites, very little association between the two ages was evident ($r = .05$). In addition, little agreement was observed between observer- and parent-ratings of anger.

Conversely, the link between observer and parent-rated measures of IC at 36 months was positive and significant. In most cases, both lab-based and parent-assessed IC were negatively associated with anger within an assessment domain (parent or laboratory). Interestingly, the lone exception to this pattern was a positive link between lab-based anger at 12 months and lab-based IC at 36 months. There was relatively little correlation between anger and IC ratings across assessments, excepting the significant negative correlation between parent anger and lab IC at 36 months. Phenotypic correlations were also calculated by sex for the three variables that showed gender differences (parent-rated and lab-assessed IC, and lab-assessed anger at 36 months). Even with our relatively large sample size, the correlation between parent and lab-based IC was the only association that showed a significant difference ($p < .05$) between males and females (the association was greater for males).

Generally, these findings indicate that children with high levels of anger have more difficulty controlling inhibitory behavior. In addition, there is relative stability in parent ratings of anger in early childhood. The lack of agreement between infant and preschool anger as assessed in the laboratory excludes genetic or environmental covariance between these two constructs. The near zero correlation also suggests that anger-related behaviors in this domain are very different from one to three years of age, and may help interpret the positive correlation between infant anger and child IC found in the laboratory.

Twin Correlations

For both observer and parent ratings of anger and IC, MZ twin intraclass correlations exceeded DZ correlations, suggesting genetic influences (Table 3). In many cases, the DZ correlation was greater than half the magnitude of the MZ correlation, indicating that shared environmental influences may be acting on these variables. Because there were mean gender differences on some of the variables, twin intraclass correlations were compared by gender and same-sex DZ twin correlations were compared to opposite-sex DZ twin correlations (results not shown in Table 3). There were no significant correlational differences by gender or DZ twinzygosity (same-sex vs. opposite-sex categories) for any of the lab-based variables. However, significant gender and DZ twin zygoty differences did occur for the questionnaire measure of IC at 36 months. The parent-assessed anger and IC cross-trait, cross-twin correlations for MZ twins exceeded those for DZ twins, suggesting that phenotypic correlations between these traits may be genetically mediated. Differences between cross-trait, cross-twin correlations for laboratory-rated anger and IC were generally smaller than the parent-rated ones, and several of the correlations were low and nonsignificant.

Genetic Model-fitting

The overall lack of covariance between parent and observer ratings in these analyses implies little genetic and environmental overlap between anger and IC across parental ratings and laboratory assessment. Therefore, multivariate genetic analyses of parent and laboratory ratings of anger and IC were conducted separately. A trivariate Cholesky model was fit for the laboratory-rated variables (infant Lab-TAB anger, preschool Lab-TAB anger, and preschool Lab-TAB IC) and the parent-assessed variables (IBQ anger, CBQ anger, and CBQ IC).

Multivariate Model-fitting Analyses: Laboratory-based Anger and IC

Table 4 presents the fit statistics for the trivariate Cholesky decomposition models for laboratory-based anger and IC. A reduced model with no shared environmental influences on 12-month anger and no genetic influences on 36-month IC (Model 2) was fit to the data based on examination of parameters from the full model and on the prior univariate model-fitting. Model 2 showed no significant decrement in fit as compared with the full model, indicating that shared environmental factors were not responsible for familial resemblance in 12-month anger, and there were no significant genetic influences on IC at 36 months of age. Model 2 was subsequently employed as a base model to test for genetic and environmental covariance. We were unable to fit reduced models without shared (Model 4) or nonshared environmental covariance (Model 5). However, genetic covariance was deleted from the model with nonsignificant change in χ^2 (Model 3). Thus, there was shared and nonshared environmental covariance between laboratory-based anger and IC, but no genetic covariance.

Significant genetic influences were noted for infant ($a^2 = .38$) and childhood ($a^2 = .32$) anger in the laboratory, but not for child IC (see Table 5). The pattern of intraclass twin correlations for IC in the laboratory suggested the presence of genetic variance; however, these twin correlations were based on complete data for each twin pair. The model-fitting analyses used all raw data, including dyads with incomplete data for one twin in the pair. The shared environment explained 23% of the variance on anger at 36 months, and 37% of the variance on IC at the same age. The remaining variance for all three variables was due to non-shared environmental influences. Genetic correlations were not estimated in the best-fitting model. Significant shared environmental covariance was present between anger and IC at 36 months ($r_c = -.73$) and was the sole contributor to phenotypic covariance between these two traits. The nonshared environmental correlation between infant anger and child IC

indicates that approximately 22% of nonshared environmental factors overlap between the variables, and nonshared environmental factors were entirely responsible for the phenotypic covariance. Of course, the lack of phenotypic association precluded finding any genetic or environmental covariance between infant and child anger.

Multivariate Model-fitting Analyses: Parent-rated Anger and IC

The fit statistics for the trivariate Cholesky decomposition models for parent-assessed anger and IC are presented in Table 6. Based on the results of the full model and univariate model-fitting, a reduced model with no shared environmental influences on 12- and 36-month anger, and 36-month IC (Model 2) was fit to the data and showed no significant change in χ^2 from the full model. Therefore, shared environmental factors were not responsible for familial resemblance in the anger and IC variables in the model and thus could not account for shared environmental covariance across the three variables. Similar to the laboratory-based analyses, Model 2 was used as a base model to test for genetic and nonshared environmental covariance. It was not possible to fit reduced models without genetic (Model 3) or nonshared environmental covariance (Model 4). Therefore, genetic and nonshared environmental covariance (but no shared environmental covariance) between parent-rated anger and IC was implicated.

Genetic influences accounted for approximately 45–72% of the variance in parent-assessed anger and IC, and the remaining variance was due to non-shared environmental influences (Table 7). Genetic correlations across the anger variables indicated that 52% of the genetic effects on anger overlapped from 12–36 months, and the genetic correlations between all anger and IC variables were negative and significant, ranging from $-.56$ to $-.26$. There was also significant nonshared environmental covariance between anger and IC at 36 months ($r_e = -.34$). Negative genetic and nonshared environmental correlations show that the same genetic and nonshared environmental factors that were associated with anger being high were also associated with IC being low. Overall, genetic factors contributed significantly to the phenotypic correlation between parent-rated anger and IC, and nonshared environmental factors contributed to this covariance at 36 months of age as well.

Discussion

This study investigated associations among laboratory and parent ratings of anger and IC in infancy and early childhood, as well as genetic and environmental influences on individual differences in these behaviors. Participants with low levels of IC typically had high levels of anger. The results of behavior genetic analyses showed significant genetic influences on parent-rated anger and IC, and on lab-assessed anger. Shared environmental variance contributed to twin similarity on lab-assessed anger and IC at 36 months. Phenotypic associations between anger and IC were largely due to genetic covariance for parent ratings, and environmental covariance in the laboratory.

The phenotypic results indicate that parent-assessed anger shows relative stability in early childhood, while lab-based anger does not. In most all cases, children with high levels of anger also had low levels of IC (lab-based anger at 12 months was positively associated with lab-based IC at 36 months). The relative lack of phenotypic associations between parent and lab-based data in this research is consistent with the historical paucity of agreement between parent and laboratory ratings of temperament, as discussed in the Introduction.

In the current study, there was no evidence of contrast effects with the parent report results. Therefore, parental bias resulting from the contrasting of DZ twins within a pair was probably not driving the divergence between modalities. Other forms of parental bias could conceivably be operating and may contribute to the phenotypic and genetic differences

between parent and lab measures. Perhaps the different assessments tap somewhat different aspects of the temperamental traits under study. In general, lab-based temperament assessment is focused on specific behaviors (e.g., angry facial expressions) in very specific tasks (e.g., escaping from the car seat), whereas parent ratings reflect a broader repertoire of behaviors across multiple broader contexts (e.g., the home, in public). Divergence between assessment modalities may also reflect differing degrees of constraint on the behavior under study. For example, children engaged in IC episodes in the laboratory were permitted to “misbehave” by eating the snack before the instructed wait period ended. At home, parents might intervene quickly when children begin to exhibit anger or fail to inhibit undesired behavior. Thus, parent perceptions may reflect a depiction of child behavior that is more externally constrained than one would observe in the laboratory.

Almost no phenotypic stability between infant and child anger was evident in the lab. The possible explanations for this lack of stability include (1) instability of individual differences in a phenotype (anger) that retains its basic nature, perhaps resulting from a change in how the phenotype was measured; (2) true developmental change in the sense that the basic nature of anger (as we measure it) changes and individual differences in the transformed variable are not preserved; or (3) unreliability of measurement that masks true stability of individual differences. Clearly, all three of the explanations can be partially accurate. Explanation (1) seems plausible because we changed how anger was assessed from 12 to 36 months, which is reasonable given the expanded response repertoire of the toddler. That is, the Lab-TAB anger episodes were more complex at 36 than at 12 months. However, when we formed the 12- and 36-month anger composites in the same way (i.e., using the same responses composited in the same manner), anger measures at 12 and 36 remained uncorrelated. Thus, explanation (1) was not supported (although it cannot be fully discounted). The second possible explanation, true developmental change, would imply that the anger measures have a different meaning at ages 12 and 36 months. This explanation holds credence (2) because we noted changed normative levels of expression and different correlates. Many more infants exhibited high levels of anger during the Lab-TAB anger episodes at 12 than at 36 months of age. Moreover, the infant anger and 36-month IC were positively correlated whereas 36-month anger and IC were negatively correlated. We therefore suggest that anger as assessed in the laboratory changes from being normative in infancy to non-normative in early childhood as children develop improved skills to control their emotions.

Associations between anger and IC help clarify the architecture of temperament by illustrating how these different aspects of child behavior are related. Typically, children with low levels of IC exhibited high levels of anger. This pattern is consistent with previous research and temperament theory (Goldsmith, Lemery, Aksan, & Buss, 2000; Gonzalez, Fuentes, Carranza, & Estevez, 2001; Gonzalez, Fuentes, & Carranza, 2003; Rothbart et al., 2001; Rueda, Posner, & Rothbart, 2005); that is, the pattern indicates that the capacity to inhibit one’s behavior overlaps with the expression of emotion in early childhood. As the child matures, IC becomes increasingly associated with planning and the ability to refrain from specific behaviors under instruction, and less related to emotion regulation. For example, Gonzalez et al. (2001) found that associations between emotionality and IC in middle childhood were significant only for females. Aksan and Kochanska (2004) have referred to an “increasing differentiation of temperament” in studies of early inhibitory behavior, and the present findings support this viewpoint by suggesting that IC is more broadly associated with emotion expression in infancy and early childhood.

Significant gender differences were apparent in parent-rated and lab-based IC and lab-assessed anger at 36 months of age. Generally, boys had higher levels of anger and lower levels of IC than girls, a finding that is consistent with previous research in this area. These

gender differences were regressed out before genetic analysis to avoid the likelihood that they would bias opposite-sex DZ correlations downward (and thus potentially bias heritability estimates upward and certainly worsen model fit). The data also showed a significant difference between male and female phenotypic correlations between parent and lab-based IC. However, this particular association was not investigated in this study via genetic analyses.

Our behavioral genetic findings confirm significant genetic influences on anger and IC as assessed by parents, and on anger as coded by observers in the lab. Interestingly, no significant genetic influences emerged for laboratory-based IC at 36 months of age. Shared environmental factors contributed to twin similarity on lab-based anger and IC at 36 months. Phenotypic covariance between anger and IC was largely due to overlapping genetic factors for parent ratings, and environmental factors in the laboratory. The presence of significant heritability for both parent-rated and laboratory-assessed anger and parent-rated IC is consistent with previous findings in early and middle childhood (Gagne & Saudino, in press; Goldsmith et al., 1997; Lemery-Chalfant et al., 2008; Deater-Deckard, et al., 2007). The lack of significant genetic variance in laboratory IC at 36 months of age is somewhat congruent with the results of the Groot et al. study (2004), in which it was unclear whether genetic or nonshared environmental variance contributed to twin similarity. It is important to note that model-fitting results did indicate that a small proportion of the variance (.17 with a confidence interval that extends to .34) in lab-based IC at 36 months was due to genetic variance; however, the parameter estimate was not significant and the reduced model (dropping genetic variance) provided a better fit.

Common family environmental factors that could influence temperament at this age include parenting effects and sharing of many experiences (e.g., same home, neighborhood, peers). Rater effects may also contribute to higher estimates of shared environmental variance for parent ratings because the same individual rates both twins. This parent rating effect would result from a shared bias across assessments (i.e., correlated error). All temperament assessments yielded substantial evidence of nonshared environmental variance, which partially may reflect measurement error. Lab-TAB coders were required to demonstrate at least 70% interrater reliability with master coders, and internal consistency for the parent assessments ranged from .61–.83. Therefore, although the nonshared environment does include measurement error, given the relative reliability of the measures, it is unlikely that nonshared environmental effects are entirely due to error. Differential parental negativity/positivity, peer groups, and child-specific experiences of life events are possible sources of nonshared environment that have been identified in previous twin studies (Plomin, DeFries, McClearn, & Rutter, 1997). If parent behavior is more or less positive toward one twin than the other, or, if one twin associates with peers who have elevated levels of behavior problems (unlikely in this case with such young participants), nonshared environmental variance could be impacted. The experience of differential negative life events could also contribute to the nonshared environment (e.g., if one child experiences an illness or an accident).

Multivariate genetic analyses of anger and IC indicated that phenotypic correlations between the two traits were largely due to common genetic factors when assessed by parents, and overlapping environmental factors when rated by observers in the laboratory. Genetic and environmental covariance was typically negative, indicating that the genetic or environmental influences that overlap contribute to elevated levels of anger and low levels of IC (and vice versa). The explanations for overlapping genetic factors include (1) pleiotropic genetic effects, whereby the same genetic influences affect multiple phenotypes, and (2) content overlap between the assessments of the two traits, which we attempted to avoid by experimental design. Overlapping environmental influences can be due to anger

and IC Lab-TAB assessments reflecting similar environmentally influenced behaviors, or shared measurement error. It is unclear why only environmental covariance is significant in the case of the Lab-TAB episodes, and genetic covariance is significant only for the parent ratings. There is no evidence of contrast effects for the parent assessments, which can bias heritability estimates upwards. The lack of phenotypic convergence between the two modalities of measurement is perhaps the more basic issue. Rather than—or in addition to—simply being a measurement problem, the lack of convergence may reflect our labeling different constructs with the same name (i.e., anger) and/or true developmental transformation of the “same” construct. It is possible that single traits are not being assessed from 12 to 36 months, within and across assessment modality. However, anger assessed in the laboratory appeared to be normative at 12 months (given the descriptive statistics reported) but non-normative at 36 months, supporting the developmental change explanation.

The behavior genetic findings for anger and IC in infancy and toddlerhood allow us to evaluate factors that contribute to individual differences in these important behavioral dimensions. The results imply that differences in IC arise from both genetic and environmental variation and thus encourage investigations to identify specific genes and specific environmental factors. The apparent exception to the pattern of genetic variance—only nonsignificant genetic variance for IC at 36 months—might reflect a decline in genetic influences on IC across development. Although environmental modifications can affect the means of even strongly heritable traits, the salience of shared and nonshared environmental influences in our study may encourage “environmental” intervention strategies when anger or IC are problematic.

The main limitation to the current research is the moderate sample size. Although sample sizes ranging from 735 to 1021 individuals for various components of the study are not small from most perspectives, the power to test certain biometric models is nevertheless limited. For example, we did not pursue formal sex limitation, rater-contrast, or sibling influence models (Neale & Cardon, 1992). Despite this drawback, the present study includes intensive, laboratory-based measures of temperament that are rarely used in projects with larger samples, which tend to rely on parental questionnaires.

Our results suggest several topics for study. At this point, a handful of investigations have indicated genetic influences on anger and IC in early childhood. However, additional developmental behavior genetic analyses that extend beyond the preschool ages are needed. Because significant genetic variance for IC was not present in this sample at 36 months, longitudinal behavioral genetic investigations of IC that extend into the school years are needed to evaluate whether heritability for IC continues to decline across development. Important gender differences in anger and IC can be better studied within a behavioral genetic framework by employing larger samples with more opposite-sex DZ twin pairs, in order to facilitate sex-limitation analyses. In addition, literature on links between these early temperament dimensions and behavior problems advocates for more twin and family studies that examine genetic and environmental covariance of temperament and behavior problems.

Acknowledgments

The Genetics of Emotional Ontogeny (GEO) project is supported by grant R37-MH50560 from the National Institute of Mental Health (P.I.: H. H. Goldsmith). The project described was also supported by Award Number T32-MH018931 from the National Institute of Mental Health (Program Director: R. J. Davidson). We also acknowledge support from P30-HD03352 and P50-MH084051. Special thanks to the students, staff, parents, and twins from GEO.

References

- Aksan N, Kochanska G. Links between systems of inhibition from infancy to preschool years. *Child Development* 2004;75:1477–1490. [PubMed: 15369526]
- Carlson SM, Davis AC, Leach JG. Less is More: Executive Function and Symbolic Representation in Preschool Children. *Psychological Science* 2005;16:609–616. [PubMed: 16102063]
- Deater-Deckard K, Petrill SA, Thompson LA. Anger/frustration, task persistence, and conduct problems in childhood: a behavioral genetic analysis. *Journal of Child Psychology and Psychiatry* 2007;48:80–87. [PubMed: 17244273]
- Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series* 1977;39:1–38.
- Derryberry D, Rothbart MK. Reactive and effortful processes in the organization of temperament. *Development and Psychopathology* 1997;9:633–652. [PubMed: 9448999]
- Eisenberg N, Cumberland A, Spinrad TL, Fabes RA, Shepard SA, Reiser M, Valiente C, Murphy SH, Losoya SH, Guthrie IK. The relations of regulation and emotionality to children's externalizing and internalizing problem behavior. *Child Development* 2001;72:1112–1134. [PubMed: 11480937]
- Eisenberg N, Spinrad TL, Fabes RA, Reiser M, Cumberland A, Shepard SA, Valiente C, Losoya SH, Guthrie IK, Thompson M. The relations of effortful control and impulsivity to children's resiliency and adjustment. *Child Development* 2004;75:25–46. [PubMed: 15015673]
- Forget-Dubois N, Perusse D, Turecki G, Girard A, Billette J, Rouleau G, Boivin M, Malo J, Tremblay RE. Diagnosing zygosity in infant twins: physical similarity, genotyping, and chorionicity. *Twin Research* 2003;6:479–485. [PubMed: 14965457]
- Gagne JR, Saudino KJ. Wait for it! The etiology of inhibitory control in early childhood. *Behavior Genetics*. (in press).
- Gartstein MA, Rothbart MK. Studying infant temperament via the revised infant behavior questionnaire. *Infant Behavior & Development* 2003;26:64–86.
- Goldsmith HH. A zygosity questionnaire for young twins: A research note. *Behavior Genetics* 1991;21:257–269. [PubMed: 1863259]
- Goldsmith HH. Studying temperament via construction of the Toddler Behavior Assessment Questionnaire. *Child Development* 1996;67:218–235. [PubMed: 8605830]
- Goldsmith HH, Buss AH, Plomin R, Rothbart MK, Thomas A, Chess S, Hinde RA, McCall RB. Roundtable: What is temperament? Four Approaches. *Child Development* 1987;58:505–529. [PubMed: 3829791]
- Goldsmith HH, Buss KA, Lemery KS. Toddler and childhood temperament: expanded content, stronger genetic evidence, new evidence for the importance of environment. *Developmental Psychology* 1997;33:891–905. [PubMed: 9383612]
- Goldsmith HH, Hewitt EC. Validity of parental report of temperament: Distinctions and needed research. *Infant Behavior & Development* 2003;26:108–111.
- Goldsmith, HH.; Lemery, KS.; Aksan, N.; Buss, KA. Temperamental substrates of personality development. In: Molfese, VJ.; Molfese, DL., editors. *Temperament and personality development across the lifespan*. Mahwah, NJ: Erlbaum; 2000. p. 1-32.
- Goldsmith, HH.; Reilly, J.; Lemery, KS.; Longley, S.; Prescott, A. The laboratory temperament assessment battery-preschool version: Description of procedures. 1995. Unpublished manuscript
- Goldsmith, HH.; Rothbart, MK. The laboratory temperament assessment battery (Technical Manual). 1999. Unpublished manuscript
- Gonzalez C, Fuentes LJ, Carranza JA. The involvement of attentional mechanisms in children's self-regulation abilities. *Cognitie Creier Comportament* 2003;7:119–132.
- Gonzalez C, Fuentes LJ, Carranza JA, Estevez AF. Temperament and attention in the self-regulation of 7-year-old children. *Personality and Individual Differences* 2001;30:931–946.
- Graham JW. Missing data analysis: Making it work in the real world. *Annual Review of Psychology* 2009;60:549–576.
- Griffin D, Gonzalez R. Correlational analysis of dyad-level data in the exchangeable case. *Psychological Bulletin* 1995;118:430–439.

- Groot AS, de Sonnevile LMJ, Stins JF, Boomsma DI. Familial influences on sustained attention and inhibition in preschoolers. *Journal of Child Psychology and Psychiatry* 2004;45:306–314. [PubMed: 14982244]
- Hwang J, Rothbart MK. Behavior genetics studies of infant temperament: Findings vary across parent-report instruments. *Infant Behavior & Development* 2003;26:112–114.
- Kochanska G, Murray K, Coy KC. IC as a contributor to conscience in childhood: From toddler to early school age. *Child Development* 1997;68:263–277. [PubMed: 9180001]
- Kochanska G, Murray KT, Harlan ET. Effortful control in early childhood: Continuity and change, antecedents, and implications for social development. *Developmental Psychology* 2000;36:220–232. [PubMed: 10749079]
- Kochanska G, Murray K, Jacques TY, Koenig AL, Vandegest KA. IC in young children and its role in emerging internalization. *Child Development* 1996;67:490–507. [PubMed: 8625724]
- Kopp CB. Antecedents of self-regulation: A developmental perspective. *Developmental Psychology* 1982;18:199–214.
- Lemery-Chalfant K, Doelger L, Goldsmith HH. Genetic relations between effortful and attentional control and symptoms of psychopathology in middle childhood. *Infant and Child Development* 2008;17:365–385.
- Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986;73:13–22.
- Mangelsdorf, SC.; Schoppe, SJ.; Buur, H. The meaning of parental reports: A contextual approach to the study of temperament and behavior problems in childhood. In: Molfese, VJ.; Molfese, DL., editors. *Temperament and personality development across the life span*. Mahwah, NJ: Lawrence Erlbaum Associates; 2000. p. 121-140.
- McGue M, Bouchard TJ Jr. Adjustment of twin data for the effects of age and sex. *Behavior Genetics* 1984;14:325–343. [PubMed: 6542356]
- Neale, M. Mx: Statistical modeling. 6. Medical College of Virginia: Department of Psychiatry; 2003.
- Neale, MC.; Cardon, LR. *Methodology for genetic studies of twins and families*. New York, NY: Kluwer Academic/Plenum Publishers; 1992.
- Neale MC, Stevenson J. Rater bias in the EASI temperament scales: A twin study. *Journal of Personality and Social Psychology* 1989;56:446–455. [PubMed: 2926639]
- O'Connor BP. SPSS and SAS programs for addressing interdependence and basic levels-of-analysis issues in psychological data. *Behavior Research Methods, Instruments, & Computers* 2004;36:17–28.
- Plomin, R.; DeFries, JC.; McClearn, GE.; Rutter, M. *Behavioral Genetics*. 3. New York: W. J. Freeman and Company; 1997.
- Plomin R, Emde R, Braungart JM, Campos J, Corley R, Fulker DW, Kagan J, Reznick S, Robinson J, Zahn-Waxler C, DeFries JC. Genetic change and continuity from 14 to 20 months: The MacArthur Longitudinal Twin Study. *Child Development* 1993;64:1354–1376. [PubMed: 8222877]
- Price TS, Freeman B, Craig I, Petrin SA, Ebersole L, Plomin R. Infant zygosity can be assigned by parental report questionnaire data. *Twin Research* 2000;3:129–133. [PubMed: 11035484]
- Reed MA, Pien DL, Rothbart MK. Inhibitory self-control in preschool children. *Merrill-Palmer Quarterly* 1984;30:131–147.
- Rothbart MK. Measurement of temperament in infancy. *Child Development* 1981;52:569–578.
- Rothbart, MK. Temperament and development. In: Kohnstamm, GA.; Bates, JA.; Rothbart, MK., editors. *Temperament in childhood*. New York: Wiley; 1989a. p. 187-247.
- Rothbart, MK. Temperament in childhood: A framework. In: Kohnstamm, GA.; Bates, JA.; Rothbart, MK., editors. *Temperament in childhood*. New York: Wiley; 1989b. p. 187-247.
- Rothbart MK, Ahadi SA. Temperament and the development of personality. *Journal of Abnormal Psychology* 1994;103:55–66. [PubMed: 8040481]
- Rothbart MK, Ahadi SA, Hershey KL, Fisher P. Investigations of temperament at three to seven years: The Child's Behavior Questionnaire. *Child Development* 2001;72:1394–1408. [PubMed: 11699677]

- Rothbart, MK.; Bates, JE. Temperament. In: Damon, W.; Eisenberg, N., editors. Handbook of child psychology. 5. Vol. 3. 1998. Social, emotional, and personality development. Hoboken, NJ: John Wiley & Sons, Inc; p. 105-176.
- Rothbart, MK.; Derryberry, D. Development of individual differences in temperament. In: Lamb, ME.; Brown, AL., editors. Advances in developmental psychology. Vol. 1. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc; 1981. p. 37-86.
- Rueda MR, Posner MI, Rothbart MK. The development of executive attention: Contributions to the emergence of self-regulation. *Developmental Neuropsychology* 2005;28:573–594. [PubMed: 16144428]
- Saudino KJ. Parent ratings of infant temperament lessons from twin studies. *Infant Behavior & Development* 2003a;26:100–107.
- Saudino KJ. The need to consider contrast effects in parent-rated temperament. *Infant Behavior & Development* 2003b;26:118–120.
- Saudino, KJ. Multiple informants. In: Everitt, BS.; Howell, DC., editors. Encyclopedia of Statistics in Behavioral Science. Chichester, UK: John Wiley & Sons, Ltd; 2005. p. 1332-1333.
- Saudino KJ, Cherny SS, Plomin R. Parent ratings of temperament in twins: Explaining the “too low” DZ correlations. *Twin Research* 2000;3:224–233. [PubMed: 11463143]
- Schmitz S, Cherny SS, Fulker DW. Increase in power through multivariate analysis. *Behavior Genetics* 1998;28:357–363. [PubMed: 9926617]
- Seifer R. Twin studies, biases of parents, and biases of researchers. *Infant Behavior & Development* 2003;26:115–117.
- Stenberg C, Campos J, Emde RN. The facial expression of anger in seven-month-old infants. *Child Development* 1983;54:178–184. [PubMed: 6831985]
- Williams BR, Ponesse JS, Schachar RJ, Logan GD, Tannock R. Development of IC across the life span. *Developmental Psychology* 1999;35:205–213. [PubMed: 9923475]
- Yamagata S, Takahashi Y, Kijima N, Maekawa H, Ono Y, Ando J. Genetic and environmental etiology of effortful control. *Twin Research and Human Genetics* 2005;8:300–306. [PubMed: 16176712]
- Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986;42:121–130. [PubMed: 3719049]

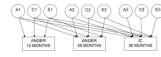


Figure 1.
Multivariate Cholesky Model

Note. A1, C1, and E1 represent the genetic, shared environmental, and nonshared environmental factors common to all three variables in the model; A2, C2, and E2 represent the genetic, shared environmental, and nonshared environmental factors common to the last two variables independent of the first; and A3, C3, and E3 are factors unique to the third variable.

Table 1

Means (Standard Deviations, and 95% Confidence Intervals) by Sex for Laboratory-and Parent-assessed Anger and IC

	Males	Females	Effect Size
Laboratory Ratings			
Anger 12 months	.05 (.95, -.04-.14)	-.05 (1.05, -.15-.05)	.10
Anger 36 months	.19 (1.01, .10-.28)	-.18 (.95, -.27- -.10)	.38*
IC 36 months	-.27 (1.06, -.36--.18)	.26 (.87, .18-.33)	-.55*
Parent Ratings			
Anger 12 months	.06 (1.0, -.04-.16)	-.07 (1.0, -.17-.03)	.13
Anger 36 months	.06 (.96, -.04-.16)	-.06 (1.04, -.17-.05)	.12
IC 36 months	-.19 (1.01, -.29--.09)	.21 (.95, .11-.31)	-.41*

Note.

* corresponding z-statistics were significant for these variables, indicating a significant sex difference.

Effect sizes estimated as Cohen's d express group differences in standard deviation units. $N=735$ children assessed by parent ratings; $N=846$ assessed for anger in the laboratory at 12 months; $N=1000$ assessed for anger in the laboratory at 36 months; and $N=1021$ children assessed for IC in the laboratory at 36 months.

Table 2
Phenotypic Correlations among Laboratory and Parent-rated Measures Within and Across Ages.

	Lab Anger 36 months	Lab IC 36 months	Parent-rated Anger 12 months	Parent-rated Anger 36 months	Parent-rated IC 36 months
Lab Anger 12 months	.03	.13**	.10*	.00	.04
Lab Anger 36 months		-.28**	-.01	.08	-.24**
Lab IC 36 months			.00	-.07	.28**
Parent-rated Anger 12 months				.34**	-.16**
Parent-rated Anger 36 months					-.45**

Note.

* $p < .05$,

** $p < .01$.

Table 3
Twin Intraclass Correlations, Cross-twin Cross-trait Correlations: Laboratory and Parent-assessed anger and IC

	1		2		3		4		5		6	
	MZ	DZ	MZ	DZ	MZ	DZ	MZ	DZ	MZ	DZ	MZ	DZ
1. Lab Anger 12 months	.38**	.17**										
2. Lab Anger 36 months	-.09	.02	.56**	.41**								
3. Lab IC 36 months	.07	-.03	-.29**	-.23**	.49**	.33**						
4. Parent-rated Anger 12 months	.06	.06	-.05	-.14*	.05	-.02	.62**	.43**				
5. Parent-rated Anger 36 months	.04	.03	.07	-.04	-.18*	-.05	.29**	.15*	.35**	.28**		
6. Parent-rated IC36 months	.01	-.05	-.14	-.12*	.17*	.04	-.22*	-.08	-.25**	-.14*	.67**	.21**

Note.

*
 $p < .05$

**
 $p < .01$.

MZ=monozygotic twins, DZ=dizygotic twins. Intraclass Correlations are bolded, Cross-twin Cross-trait Correlations are under the diagonal.

Table 4

Fit Statistics for Multivariate Models of Laboratory-assessed 12-month Anger, and 36-month Anger and IC

	Overall Fit of Model			Relative Fit of Model		
	-2LL	df	AIC	$\Delta\chi^2$	Δdf	p
1. Full Model	7716.42	2837	2042.42			
2. No C on Anger 12, No A on IC 36	7718.22	2843	2032.22	1.80	6	.94
3. Model 2 plus No A Covariance	7718.49	2844	2030.49	2.07	7	.96
4. Model 2 plus No C Covariance	7760.01	2844	2072.01	43.58	7	.00
5. Model 2 plus No E Covariance	8125.21	2846	2433.21	408.79	9	.00

Note. -2LL = Likelihood Statistic. *df* = Degrees of freedom. AIC = Akaike's Information Criterion. $\Delta\chi^2$ = Difference in Chi-square between the full model and reduced models. Δdf = Difference in *df* between the full model and reduced models. Best-fitting model in bold. *N* = 846 assessed for anger in the laboratory at 12 months; *N* = 1000 assessed for anger in the laboratory at 36 months; and *N* = 1021 children assessed for IC in the laboratory at 36 months.

Table 5

Multivariate Estimates of Genetic and Environmental Variance, and Genetic and Environmental Correlations (and 95% Confidence Intervals) for Laboratory-assessed 12-month Anger, 36-month Anger and 36-month IC using the best-fitting Model

	Variance Components			Genetic and Environmental Correlations		
	a^2	c^2	e^2	r_g	r_c	r_e
Anger 12 months	.38 (.26-.49)	-	.62 (.51-.75)			
Anger 36 months	.32 (.05-.53)	.23 (.08-.42)	.45 (.36-.56)			
IC 36 months	-	.37 (.29-.44)	.63 (.56-.71)			
Anger 12-Anger 36				-	-	.07 (-.05-.18)
Anger 12-IC 36				-	-	.22 (.11-.31)
Anger 36-IC 36				-	-.73 (-1.0- -.47)	-.05 (-.15-.06)

Note. a^2 = % genetic variance, c^2 = % shared environmental variance, e^2 = % nonshared environmental variance, r_g = genetic correlation, r_c = shared environmental correlation, r_e = nonshared environmental correlation.

Table 6

Fit Statistics for Multivariate Models of Parent-assessed 12-month Anger, 36-month Anger and 36-month IC

	Overall Fit of Model			Relative Fit of Model		
	-2LL	df	AIC	$\Delta\chi^2$	Δdf	p
1. Full Model	5756.28	2175	1406.28			
2. No c on Anger 12, 36, and IC 36	5759.65	2181	1397.65	3.37	6	.76
3. Model 2 plus no A covariance	5812.98	2184	1444.98	56.71	9	.00
4. Model 2 plus no E covariance	5785.55	2184	1417.55	29.27	9	.00

Note. -2LL = Likelihood Statistic. *df* = Degrees of freedom. AIC = Akaike's Information Criterion. $\Delta\chi^2$ = Difference in Chi-square between the full model and reduced models. Δdf = Difference in *df* between the full model and reduced models. Best-fitting model in bold. *N*=735.

Table 7

Multivariate Estimates of Genetic and Environmental Variance, and Genetic and Environmental Correlations (and 95% Confidence Intervals) for Parent-assessed 12-month Anger, 36-month Anger and 36-month IC using the best-fitting Model

	Variance Components			Genetic and Environmental Correlations		
	a^2	c^2	e^2	r_g	r_c	r_e
Anger 12 months	.72 (.64-.79)	-	.28 (.21-.36)			
Anger 36 months	.45 (.32-.55)	-	.56 (.45-.68)			
IC 36 months	.63 (.51-.72)	-	.37 (.28-.49)			
Anger 12-Anger 36				.52 (.37-.66)	-	.14 (-.02-.29)
Anger 12-IC 36				-.26(.32-.62)	-	.07 (-.10-.24)
Anger 36-IC 36				-.56(-.70- -.40)	-	-.34 (-.47- -.19)

Note. a^2 = % genetic variance, c^2 = % shared environmental variance, e^2 = % nonshared environmental variance, r_g = genetic correlation, r_c = shared environmental correlation, r_e = nonshared environmental correlation.