Behav Genet (2010) 40:327–337
DOI 10.1007/s10519-009-9316-6

ORIGINAL RESEARCH

Wait For It! A Twin Study of Inhibitory Control in Early Childhood

Jeffrey R. Gagne · Kimberly J. Saudino

Received: 3 June 2009 / Accepted: 7 November 2009 / Published online: 21 November 2009

Abstract Inhibitory control (IC) is a dimension of child temperament that emerges in toddlerhood and involves the ability to regulate behavior in response to instructions or expectations. In general, children with low levels of IC have more cognitive and social difficulties, and higher levels of problem behaviors. Unfortunately, there is a paucity of research on the heritability of this important behavioral dimension. The present study used a twin design to examine the extent to which genetic and environmental factors contribute to individual differences in IC. Laboratory and parent assessments of IC were conducted on 294 same-sex twin pairs (133 MZ, 161 DZ) at 24 months of age. Model-fitting analyses showed that genetic factors accounted for 38 and 58% of the variance in laboratory- and parent-rated IC, respectively. Multivariate genetic analyses also revealed that the covariance between observed and parent-assessed IC could be predominantly explained by common genetic influences.

Keywords Temperament · Inhibitory control · Childhood · Twins · Laboratory ratings

Introduction

Inhibitory control (IC) is a dimension of temperament that develops over early childhood and involves the regulation of behavior, typically in response to instructions or expectations. Individuals who develop sufficient IC are able to inhibit behavior when it is necessary. Deficient or underdeveloped IC often results in impulsive behavior. In general, children with typically developing IC have fewer cognitive difficulties, more stable temperaments, and less behavioral maladjustment (Eisenberg et al. 2001, 2004; Hughes et al. 2000; Schachar et al. 1995; Wolfe and Bell 2003).

Child temperament is defined as individual differences in both reactivity and self-regulation, is moderately stable, and collectively forms the basis of later personality (Derryberry and Rothbart 1997; Goldsmith et al. 1987; Rothbart and Ahadi 1994; Rothbart and Bates 2006). In the Rothbart model of temperament, IC is the latest developing component of the broad effortful control (EC) factor. EC emerges in infancy and the second year, is considered self-regulatory, and involves the efficiency of executive attention, including the ability to inhibit a dominant response (IC) and the ability to activate responses (Derryberry and Rothbart 1997; Rothbart 1989a, b; Rothbart and Ahadi 1994; Rothbart and Bates 2006). IC becomes distinct as a dimension of temperament around the age of 2 years, and continues to develop in the toddler and preschool years (Kochanska et al. 1996; Rothbart 1989a). It is important to note that although IC and EC are intended as overlapping yet somewhat distinct constructs (Rothbart and Bates 2006) the terms are sometimes used interchangeably to describe the same behaviors.

IC is positively associated with executive functioning skills such as working memory, planning and attentional flexibility (Chiappe et al. 2000; Conway et al. 2000; Pallodino et al. 2003; Passolunghi and Siegel 2001; Wolfe and Bell 2003), as well as theory of mind (Frye et al. 1995; Perner et al. 1987; Perner et al. 1999; Russell et al. 1991).

Edited by Deborah Finkel.

J. R. Gagne (✉)
Department of Psychology, University of Wisconsin-Madison, 1202 West Johnson Street, Madison, WI 53706-1611, USA
e-mail: jgagne@wisc.edu

K. J. Saudino
Boston University, Boston, MA, USA

Children who develop appropriate levels of IC also have increased levels of social competence in middle childhood (Lengua 2003; Nigg et al. 1999; Olson et al. 1999). Although these social and cognitive outcomes are consistently linked to IC, the importance of investigating the etiology of IC is largely attributed to relations with child behavior problems and psychopathology, and the potential role of IC as an endophenotype for relevant childhood disorders.

The behavioral disorder that is most strongly associated with IC in childhood is Attention Deficit Hyperactivity Disorder (ADHD; Pliszka et al. 1997; Schachar and Logan 1990; Schachar et al. 1995). ADHD-diagnosed children typically have poorer IC as compared to children with diagnoses of anxiety disorders and oppositional defiant disorder (ODD), who in contrast to ADHD children do not typically show deficits in IC (Oosterlaan et al. 1998; Oosterlaan and Sergeant 1996; Schachar et al. 2000). However, lower levels of IC may lead to an increased risk for other forms of psychopathology, including disorders co-morbid with an ADHD diagnosis in childhood. In fact, in at least one meta-analysis, children with conduct disorder (CD) display similar levels of IC deficits as those with ADHD, or those co-morbid for ADHD and CD (Oosterlaan et al. 1998). In addition, similar decrements in IC were observed in children with co-morbid ADHD and ODD in one other study (Scheres et al. 2001).

Although a large body of research supports the presence of an inhibition deficit in ADHD, several more recent investigations suggest that inhibitory deficits may not be homogenous. A meta-analysis of inhibitory ability and ADHD indicates that the Stroop Color-Word Task does not support an "interference control" deficit in 6–27 year-olds with ADHD (Van Mourik et al. 2005). The authors posit that inhibition deficits in ADHD are largely the result of inappropriate methods used to calculate interference scores with the Stroop task. Another review supports the theory that impaired executive functioning in ADHD is not homogenous, but rather, involves multiple cognitive profiles that include poor IC as indexed by both standard response inhibition *and* delay aversion tasks (Castellanos et al. 2006). Performance on these two different tasks in ADHD children are considered dissociable processes and may reflect differing patterns of underlying cognitive dysfunction. In a twin study of attention problems and IC assessed with the Stroop task, associations between these phenotypes were non-significant after controlling for IQ (Polderman et al. 2009). The relation between IC and attention problems resulted from genetic factors that are shared between IC and IQ.

These more recent conflicting findings coupled with the neuropsychological heterogeneity of children with ADHD (Castellanos et al. 2006) temper some of the previous enthusiasm for a specific IC or inhibitory deficit in children with ADHD. One obvious explanation for varied results lies in the differing assessments of IC in studies of ADHD. The use of Stroop tasks has been called into question, however, response inhibition (i.e., the stop-signal task) and delay aversion tasks may be more appropriate. Another issue is the wide range of ages that have been investigated and summarized in these studies. A very small minority of investigations have examined IC at the level of etiology in early childhood—the developmental period when IC is thought to emerge. The lack of research in this area is driven by the fact that neurocognitive measures such as the standard stop-signal and Stroop tasks are typically inappropriate for very young children. Most studies approach IC as a fully developed executive function rather than as an aspect of early temperament. In the present research, we focus on IC as an early emerging dimension of temperament as assessed with delay ability tasks.

Although no prior behavioral genetic studies have examined individual differences in IC in early childhood, twin studies on executive functioning measures of IC in older children (Groot et al. 2004; Polderman et al. 2009) suggest that individual differences in IC at age two *might* be explained by genetic factors. Six-year-olds tested in a Go-NoGo task that assessed IC-like behaviors (237 twin pairs) displayed familial resemblance, but model-fitting analyses could not distinguish between genetic or shared environmental effects (Groot et al. 2004). In a twin study of 9-, 12-, and 18-year-old-twins using a Stroop task to assess IC, genetic variance ranged from 39 to 51% across age (Polderman et al. 2009). Similar behavioral genetic investigations find genetic variance present in EC, but the focus was on a broader construct (EC) in older children and adults using parent- and self-assessments (Goldsmith et al. 1997; Lemery-Chalfant et al. 2008; Yamagata et al. 2005).

Further research is required to clarify genetic and environmental contributions to individual differences in IC. In the present study we examined the etiology of IC in 2-year-old twins, the age at which IC emerges as an important dimension of temperament. Because parent ratings of temperament may be prone to rater biases that exaggerate differences between DZ twins and can lead to overestimates of heritability (Neale and Stevenson 1989; Plomin et al. 1993; Saudino 2003), we used both parent ratings and lab-based observer ratings to provide a more accurate picture of the genetic and environmental influences on IC. In addition to examining genetic and environmental influences on IC as assessed by each method, we also explore genetic and environmental influences on the covariance between the two methods. As predicted by previous investigations, we expected that IC would be, in part, influenced by genetic factors. We also expected that both methods of IC would be associated and would have

some genetic influences in common. More conclusive findings will add to our current knowledge of the etiology of IC, which has not been extensively studied in twin samples using lab- and parent-based assessments. Acknowledging that the nature of IC is complex and different components emerge at different points in development, this study focuses on the first emerging component of IC, ability to delay or simple response inhibition (vs. complex response inhibition; see Garon et al. 2008).

Method

Participants

Families were recruited from a database of names and addresses of twin births supplied by the Massachusetts Registry of Vital Records (screened for infant mortality). Over 40% of the families contacted agreed to participate in the study. Participants included 294 same-sex twin pairs (133 MZ, 161 DZ). The twins were assessed at 24 months of age in a laboratory situation (mean age = 2.07 years, SD = .05). The sample was derived from the Boston University Twin Project (BUTP), a study of child temperament with a focus on the multi-method assessment of activity level and related behaviors in early childhood. Zygosity was assessed through DNA cheek swab samples using 10 multiplex markers. All twins were screened for birth weight (more than 1,750 g), gestational age (over 34 weeks), and no significant health problem that may interfere with behavioral testing (e.g., cerebral palsy).

The sample included 157 male and 137 female twin pairs in this study which were approximately equally distributed across zygosity (51% male MZ; 55% male DZ). The racial composition of the sample was 88.2% White, 3.1% Black, 2.1% Asian, and 6.6% mixed. Although the majority of the participants were White, this sample is representative of the state of Massachusetts (i.e., 86.5% White, 6.9% Black, 4.9% Asian, 1.3% mixed; 2006 Census of Population and Housing). The average socioeconomic status (SES) of the twins was predominantly middle class according to the Hollingshead index (mean SES = 51.2, SD = 10.87), although there was considerable range (22–66) in the sample.

Procedure

The procedure involved two visits, 48 h apart, to the BUTP laboratory. With the exception of two families, all participants returned within 2 days for the second visit (99%) and the average time in hours between visits was 48.2 (SD = 3.2). During the first visit, one twin was assessed in a standardized test setting while the other twin was assessed in a laboratory play situation. While the twins were participating in the assessments, parents completed consent forms and were provided with DNA swab kits and questionnaires about each child's temperament and family demographics to be completed for the next visit. During the second visit, the twins underwent the assessments that their co-twin participated in the first day, and parents returned all questionnaires and DNA kits. All assessments were conducted by trained testers (within a twin pair, each twin was assessed by a different tester). In the laboratory play situation, twins participated in episodes from the Laboratory Temperament Assessment Battery—Preschool Version (Lab-TAB; Goldsmith et al. 1995), a comprehensive laboratory-based temperament assessment that includes behavioral episodes corresponding to specific dimensions of temperament. The play situation typically took less than 1 h to complete. The assignment of first- and second-born twins to play and test situations was counterbalanced across the study such that half of all first-born twins participated in the play situation at the first lab visit and half in the second lab visit.

Observer ratings of inhibitory control

Observed IC was assessed using the Lab-TAB IC episodes, which include "Dinky Toys," "Snack Delay," and "Gift." These are categorized as delay ability tasks and were adapted from Kochanska's work on IC and EC (Kochanska et al. 1996, 2000). All Lab-TAB coders were trained by master coders and required a 90% inter-rater reliability criterion before they were permitted to code episodes on their own. Ten percent of the sample was rated by a second observer and the interrater agreement for the Lab-TAB IC composite was high ($r = .89$, $p < .01$).

In the Dinky Toys episode, the child was asked to select one out of an array of six attractive trinkets, thereby inhibiting the urge to pick more than one toy or hoard all of them. There were two separate trials included in this episode. Both trials were coded for the child's initial approach to the stimuli, latency to touch the first toy, latency to choose a toy, style of touching, frequency of touch, number of toys touched, level of distress, following directions, comprehension and interest, and a global rating of impulsivity.

During Snack Delay, the child was offered a snack (a candy or a cracker), 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 retrieve 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. The child's global IC was coded during each trial, as well as during the instruction phase when the experimenter

reminded the child of the rules of the game. In addition, each trial was coded for whether the child waited to eat the snack, the presence of fidgeting and self-distracting behaviors, latency to fidget and self-distract, and latency to eat the snack.

The Gift episode involved the inhibition of the desire to open a gift presented to the child. The child was given a small wrapped gift and was required to wait 2 min before opening it. Coding for Gift consisted of IC during instruction phase, fidgeting and distraction behavior across the 2 min, latency to fidget, latency to self-distract, latency to open the gift, frequency of self-corrections, distress, and whether or not the child opened the gift before the end of the 2-minute waiting interval.

A composite IC score was formed following guidelines from Goldsmith et al. (1995). The variables used in the summary scores for each episode were selected on the basis of principal component analyses (not all variables assessed in each episode were used in the episode composites). All item-level data were converted to z-scores and averaged across trials. The Dinky Toys summary score was created by averaging style of touching, frequency of touch, number of toys touched, latency to choose, following directions, and impulsivity scores. The factor loadings for the Dinky Toys principal components analysis ranged from .59 to .88. The Snack Delay summary score was based on the mean of global IC, whether the child waited to eat the snack, and latency to eat (factor loadings ranged from .71 to .93). The summary score for Gift comprised opening the gift early, frequency and latency to self-distract, and latency to open the gift (factor loadings ranged from .75 to .85). Participants with low IC did not self-distract. A small number of variables that are conceptualized as pertaining to IC were also eliminated through the results of principal component analyses (e.g., IC during instruction phase for Gift). All Lab-TAB summary scores were significantly correlated, and an overall composite of observed IC was computed from the mean of these summary scores (after converting to z-scores).

Parent ratings of inhibitory control

The Lab-TAB ratings of IC were complemented by parent ratings of IC on the Toddler Behavior Assessment Questionnaire-Revised (TBAQ-R; Goldsmith 1996). Questionnaires were mainly completed by mothers, with a small proportion completed by fathers (94 and 6%, respectively). The TBAQ-R requires the parent to make judgments of behaviors in specific situations observed within the past month (e.g., "When asked to wait for something (like a toy or a snack), how often did your child find something to distract her/himself until it was time?") and is rated on a scale from 1 (*never*) to 7 (*always*). This measure is well-established, reliable, valid, and is appropriate for children 2 years of age (Goldsmith 1996). Published estimates of internal consistencies for the TBAQ-R range from .86 to .89 (Goldsmith 1996), and in the current study, Cronbach's alpha was .82 for the IC subscale.

Descriptive statistics and phenotypic correlations

Descriptive statistics, tests of mean sex and zygosity differences, and phenotypic correlational analyses were conducted for parent- and observer-rated IC. To account for the nested nature of twin data we used Generalized Estimating Equation models to test for mean differences (sex and zygosity) in IC (Liang and Zeger 1986; Zeger and Liang 1986), and dyad-level correlations were computed following procedures outlined by Griffin and Gonzalez (Griffin and Gonzalez 1995; O'Connor 2004).

Twin correlations

Intraclass twin correlations provided an index of twin similarity for the trait under study. If MZ intraclass twin correlations exceed DZ intraclass twin correlations for a trait, it suggests that genetic factors contribute to individual differences for that trait. The foundation of multivariate approaches to behavioral genetics is the cross-method, cross-twin correlation whereby twin A's score on one variable (e.g., observed IC) is correlated with twin B's score on the other variable (e.g., parent-assessed IC), and vice versa. When cross-method cross-twin correlations for MZ twins exceed those of DZ twins, it suggests genetic contributions to the covariance between variables. 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 and Bouchard 1984).

Univariate model-fitting analyses

To estimate genetic and environmental variance components and provide a more accurate test of genetic and environmental effects univariate behavioral genetic models were fit to twin variance/covariance matrices using Mx maximum-likelihood model-fitting procedures (Neale et al. 2003). Univariate model-fitting began with testing the full ACE model and model fit was assessed by the χ^2 goodness-of-fit test. Alternate models were tested and compared to the full model. These included reduced models with shared environmental effects dropped (AE), additive genetic effects dropped (CE), and both dropped (E). The AE model assumes that all familial resemblance is due to genes; the CE model assumes resemblance is due to shared environment; and the E model assumes no familial resemblance.

Reduced models were assessed with the χ^2 difference test to determine if they were significantly different from the full model. Heritability estimates, environmental variances, and their 95% confidence intervals were estimated using the best-fitting model.

Multivariate model-fitting analyses

Multivariate model-fitting analyses were used to examine genetic and environmental contributions to the covariance between parent and laboratory ratings of IC. A Cholesky decomposition model was fit to observed covariance matrices (Fig. 1). The latent variables A1, C1, and E1 represent the overlapping genetic, shared environmental, and nonshared environmental factors that influence both methods of assessing IC; and A2, C2, and E2 are factors unique to parent-rated IC.

This model allows the estimation of genetic and environmental correlations (i.e., r_g, r_c, r_e) between the two assessments of IC (i.e., the degree to which genetic or environmental factors for one variable overlap with those on another, *independent of the heritability* of each variable); and the genetic and environmental contributions to phenotypic correlations between phenotypes (i.e., the extent that these overlapping influences account for the phenotypic correlation between two traits). Alternate models were tested and compared to the full model using the χ^2 difference test. Specifically, the A and C variances for each variable, and A, C, and E covariances between variables were eliminated from the model to determine if genetic and/or environmental variances and covariances were significant.

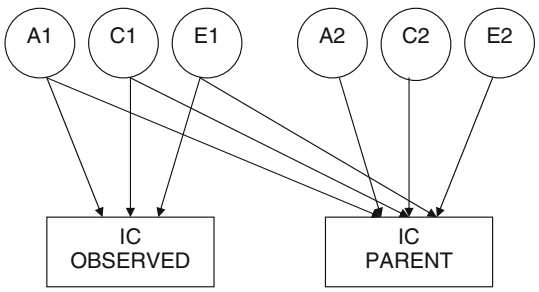

Fig. 1 Bivariate cholesky model

Results

Descriptive statistics

Table 1 lists the means and standard deviations of observer- and parent-rated IC for both males and females across twin zygosity. There were no significant mean differences between MZ and DZ twins for the two IC measures (observer: $z = -.56$, $p = .57$; parent-rated: $z = -.07$, $p = .95$). For both measures, males were significantly lower in IC than females (Lab-TAB: $z = 2.57$, $p < .01$; TBAQ: $z = 4.01$, $p < .01$). The effect sizes as measured by Cohen's d indicated that the mean for males was approximately 24% of a standard deviation lower than females on observed IC, and 43% of a standard deviation lower on parent-assessed IC. Agreement between observer- and parent-rated IC was relatively modest ($r = .21$, $p < .01$).

Genetic and environmental influences on individual differences in inhibitory control

For both observer- and parent-rated IC, MZ twin intraclass correlations (Table 1) were higher than DZ correlations, suggesting genetic influences. With the exception of the DZ twin correlation for observer-rated IC, all twin intraclass correlations were significant at the $p < .01$ level. Differences between MZ and DZ intraclass correlations were also significant ($p < .05$ for observed IC, $p < .01$ for parent-assessed IC). The DZ correlation for parent-rated IC was greater than half the magnitude of the MZ correlation and hints that shared environmental influences may also be influencing this variable.

Univariate model fitting results for observer- and parent-rated IC are presented in Table 2. For observed IC on the Lab-TAB, the full ACE model fit the data well. Reduced models with no genetic variance (i.e., CE model) or no

Table 1 Means (and standard deviations) by sex and zygosity, effect sizes of sex and zygosity differences, and twin intraclass correlations for observed and parent-rated inhibitory control

	Males		Females		Effect size		Twin intraclass correlations	
	MZ ($n = 136$)	DZ ($n = 178$)	MZ ($n = 130$)	DZ ($n = 144$)	Sex	Zyg.	MZ ($n = 266$)	DZ ($n = 322$)
Observed IC	−.02 (.63)	−.10 (.69)	.08 (.65)	.10 (.65)	−.24	.05	.38	.16
Parent-rated IC	37.06 (9.17)	37.27 (8.82)	40.92 (8.19)	40.87 (8.64)	−.43	.01	.84	.55

Note: *MZ* monozygotic twins, *DZ* dizygotic twins, *Zyg*. zygosity. Effect size estimated as Cohen's *d* express group differences in standard deviation units

Table 2 Univariate estimates of genetic and environmental variance (and 95% CI) and model fit statistics for observed and parent-rated inhibitory control

	Variance estimates			Overall fit of model				Relative fit of model[a]		
	h^2	c^2	e^2	χ^2	*df*	*p*	AIC	$\chi^{2\text{diff}}$	df^{diff}	*p*
Observed IC models										
ACE	.38 (.06–.51)	.00 (.00–.23)	.62 (.49–.77)	1.35	3	.72	−4.65	–	–	–
AE	**.38 (.24–.51)**	**–**	**.62 (.49–.76)**	**1.35**	**4**	**.85**	**−6.65**	**0.00**	**1**	**1.0**
CE	–	.25 (.14–.36)	.75 (.64–.86)	6.46	4	.17	−1.54	5.11	1	.02
E	–	–	1.0	25.44	5	.00	15.44	24.09	2	.00
Parent-rated IC models										
ACE	**.58 (.39–.82)**	**.26 (.03–.44)**	**.16 (.12–.21)**	**2.59**	**3**	**.46**	**−3.41**	**–**	**–**	**–**
AE	.84 (.80–.88)	–	.16 (.12–.20)	7.51	4	.11	−0.49	4.92	1	.03
CE	–	.68 (.62–.74)	.32 (.26–.38)	41.05	4	.00	33.05	38.46	1	.00
E	–	–	1.0	224.79	5	.00	214.79	222.20	2	.00

Note: h^2 heritability; c^2 shared environmental variance; e^2 nonshared environmental variance; χ^2 Chi-square fit statistic; *df* degrees of freedom; *AIC* Akaike's Information Criterion (AIC is used to compare the relative fit of models, those with a lower AIC are judged to fit better); $\chi^{2\text{ diff}}$ Chi-square difference between full ACE model and reduced model; df^{diff} *df* difference between full ACE model and reduced model; *A* additive genetic effects; *C* shared environmental effects; *E* nonshared environmental effects. Boldface denotes best fitting model

[a] Relative fit of the model determined by the $\chi^{2\text{diff}}$ between full ACE model and reduced model

familial resemblance (i.e., E model) resulted in significant decrements in fit. However, a reduced model with no shared environmental influences (i.e., AE model) did not result in a significant change in χ^2, indicating that shared environmental factors were not present and that genetic factors were responsible for the familial resemblance in observer-rated IC. Genetic influences accounted for approximately 38% of the variance in observed IC and the remaining variance was due to non-shared environmental influences.

For parent-rated IC on the TBAQ, a different pattern of results emerged. In this case, the full ACE model fit the data well, and all reduced models (i.e., dropping genetic or shared environmental variances) provided significantly worse fits to the data. Therefore, both genetic and shared environmental factors explained twin resemblance for parent-rated IC. Genetic factors accounted for 58% of the variance in parent-rated IC, shared environmental factors explained 26% of the variance, and non-shared environmental factors accounted for the remaining variance.

To what extent do genetic influences on parent-rated IC overlap with genetic influences on observer-rated inhibitory control?

The observer- and parent-rated IC cross-method cross-twin correlation for MZ twins ($r = .18$, $p < .05$) exceeded that for DZ twins ($r = .12$, ns). With such low correlations and our relatively small twin sample, the difference between MZ and DZ cross correlations was not significant. Nevertheless, the overall pattern hints that the phenotypic correlation between observer- and parent-assessed IC may be genetically mediated. This highlights the need for our more powerful multivariate genetic model-fitting analyses (see Table 3). Although not significant in the univariate model for observer-rated IC, the C parameter was retained in the bivariate model to allow for a more powerful test of possible shared environmental influences. The full model (i.e., ACE observed IC/ACE parent-rated IC) provided a good fit to the data. Dropping shared environmental influences for observed IC (i.e., AE/ACE) showed no significant decrement in fit, confirming that shared environmental variance is not significant for observed IC and consequently, there can be no significant shared environmental covariance between observed and parent-rated IC. This model with genetic and nonshared environmental influences on observed IC and genetic, shared and nonshared environmental influences on parent-rated IC was used as a base model to test for genetic and nonshared environmental covariances.

As indicated in Table 3, it was possible to fit a reduced model with no nonshared environmental covariance (i.e., AE/ACE No E Covariance), but not one without genetic covariance between the two measures of IC (i.e., AE/ACE No A Covariance). Therefore, there is significant genetic covariance between observer- and parent-rated IC, but no significant environmental covariance. The genetic correlation indicates that 47% of the genetic effects on observed IC overlap with genetic effects on parent-rated IC (Table 3). Moreover, it is only these genetic factors that contribute significantly to the phenotypic correlation between observer- and parent-rated IC. Because the initial

Table 3 Fit statistics for bivariate models of observed and parent-rated inhibitory control, and variance estimates and genetic and environmental correlations for the best-fitting bivariate model of observed and parent-rated inhibitory control (95% CI)

Fit statistics for bivariate models

	Overall fit of model				Relative fit of model[a]		
	χ^2	*df*	*p*	AIC	χ^{2diff}	df^{diff}	*p*
1. ACE Observed/ACE Parent-rated	11.02	11	.44	−10.98	–	–	–
2. AE Observed/ACE Parent-rated	11.06	13	.61	−14.94	.04	2	.98
3. Model 2 plus No A Covariance	32.09	14	.00	4.09	21.07	3	.00
4. Model 2 plus No E Covariance	**11.13**	**14**	**.68**	**−16.87**	**.11**	**3**	**.99**

Variance estimates and genetic and environmental correlations (best-fitting model)

	Variance estimates			Genetic and environmental correlations		
	h^2	c^2	e^2	r_g	r_c	r_e
Observed IC	.39 (.25–.51)	–	.61 (.49–.75)			
Parent-rated IC	.58 (.39–.80)	.27 (.05–.44)	.15 (.12–.20)			
Observed IC-Parent-rated IC				.47 (.28–.68)	–	–

Note: χ^2 Chi-square fit statistic; *df* degrees of freedom; *AIC* Akaike's Information Criterion (AIC is used to compare the relative fit of models, those with a lower AIC are judged to fit better); χ^{2diff} Chi-square difference between full ACE/ACE model and reduced model; df^{diff} *df* difference between full ACE/ACE model and reduced model; *A* additive genetic effects; *C* shared environmental effects; *E* nonshared environmental effects; h^2 heritability; c^2 shared environmental variance; e^2 nonshared environmental variance; r_g genetic correlation; r_c shared environmental correlation; r_e nonshared environmental correlation. Boldface denotes best fitting model. Shared and nonshared environmental correlations were not included in the best-fitting model

[a] Relative fit of the model determined by the χ^{2diff} between full ACE model and reduced model

phenotypic correlation was somewhat small, it is important to mention that this shared genetic variance between the IC assessments is also relatively small.

Discussion

This study is the first to examine genetic and environmental influences on individual differences in IC in early childhood using both parent and laboratory assessments. Unfortunately, most past twin research on temperament has employed global parent ratings of behavior only. An extensive literature indicates a lack of convergence between parent and laboratory ratings of temperament (Goldsmith et al. 1991; Mangelsdorf et al. 2000; Saudino and Cherny 2001; Saudino et al. 2004; Seifer et al. 1994), suggesting that these modes of assessment tap different aspects of child temperament. A recent BUTP investigation found that parent ratings of activity level tap different genetic and environmental factors than observer and mechanical ratings of activity level (Saudino 2009). Using both parent and observer ratings may provide a more accurate estimation of genetic and environmental factors that influence IC than relying on parent ratings alone. Examining both parent and observer assessment of IC also allows for the investigation of covariance between the two types of assessment.

Results confirm that IC is both genetically and environmentally influenced (parent-rated assessments of IC showed shared environmental influences). Consistent with previous temperament research, agreement between parent and laboratory ratings was modest. However, there is moderate overlap between the genetic factors that influence observed IC and those that influence parent-assessed IC. Although the genetic correlation is moderate, it is these overlapping genetic effects that entirely explain the phenotypic correlation. The presence of significant heritability for both parent-rated and laboratory-assessed IC provides initial support for genetic influences on IC in early childhood, and is consistent with the Polderman et al. (2009) findings in older twins (9–18 years of age). These results contrast to those of Groot et al. (2004), who found significant familial resemblance but could not differentiate between genetic and shared environmental variance in a study with a similar sample size employing 6-year-old twins.

In this investigation, IC is examined as it emerges at 2 years of age as opposed to the Groot et al. (2004) paper and other previous twin studies that focus on IC in older children. The use of simple delay ability tasks in the lab and a broadly self-regulatory parent assessment of IC may provide an advantage for detecting an effect over the complex response inhibition tasks that were used in the Groot study. Correspondingly, origins of individual

differences in the more complex and later emerging aspects of IC such as performance on response inhibition and Stroop tasks in early school age through adulthood may differ from the tasks used in this study. Another possibility is that the heritability of IC may decrease after toddlerhood until later school age. Groot et al. (2004) did not find significant genetic influences on IC in 6-year-olds using laboratory measures, yet we find genetic influences using both parent and lab ratings in toddlers. The results of the Polderman et al. (2009) study indicate increases in heritability from the ages of 9–18 years. While it is possible that heritability could decrease from 2 to 6 and increase from 9 to 18, we would suggest that it is more likely that methodological factors explain the discrepant findings for IC. Longitudinal analyses of IC that extend from early to middle childhood are needed to explore the question of developmental changes in genetic and environmental influences on IC. To that end, we plan to follow-up the sample at later ages.

Shared environmental influences were significant for only parent-rated IC, consistent with previous research showing shared environmental influences on the temperament dimensions of anger proneness (Deater-Deckard et al. 2007; Goldsmith et al. 1997), and positive affect (Goldsmith et al. 1999). Common family environmental factors that could influence parent-rated IC include parent personality, and twins sharing many of the same experiences (e.g., same home, neighborhood, playmates). Previous research indicates that positive parent personality and parent–child relations that reflect these traits enhance childhood IC (Kochanska et al. 2000; Olson et al. 1990). Parent personality would be a shared environmental influence because the parent's personality is a constant across twins (note this is not the same as parental responsiveness to twins' behaviors), so it is possible that the parent's personality has a common effect on the development of IC for both twins. However, if parent personality were to operate in this manner one would expect to also find shared environmental influences on the observed measure of IC as well as our parent-rating measure. The same logic would apply to parenting behaviors. It is possible that shared environmental effects are less apparent in the laboratory because much of the behavior in the Lab-TAB activities is child-directed, whereas home IC behavior may be more parent-directed. Rating effects may also contribute to higher estimates of shared environmental variance because the same parent rates both twins, thus rater biases, response styles (e.g., the tendency to rate leniently or severely), normative standards, and implicit theories (e.g., stereotypes, halo effects) could contribute to shared environmental variance (Saudino 2005a). These parent rating effects may contribute to the differences in findings between parent and observer ratings of IC.

Both IC assessments yielded substantial evidence of nonshared environmental variance, and estimates for observed IC were larger than those for parent-rated IC. This higher estimate of nonshared environmental variance for the Lab-TAB may reflect greater measurement error. This does not appear to be the case. Both methods of assessing IC demonstrated high reliability suggesting that although nonshared environment does include measurement error, it is unlikely that nonshared environment is entirely due to error. Moreover, there was no evidence to suggest that observer ratings were less reliable than parent ratings on the TBAQ. This raises more interesting possibilities for sources of nonshared environment that are specific to each child. Differential parental negativity/positivity, assessment context effects, and child-specific experiences of life events are possible sources of nonshared environment that may play a role in the etiology of IC. As previously mentioned, parent positivity and positive parent–child relationships are associated with IC development in childhood. If parents differ in their positivity towards each twin, nonshared environmental variance in IC could be impacted. Differences between twins could also arise from context effects in the lab, whereby aspects of the lab situation are not consistent across the twins (e.g., different testers or days of assessment). Contextual differences in the lab situation reflect specific testing effects as opposed to anonymous measurement error. For example, the Lab-TAB tasks, stimuli and procedures were the same for each twin, but twins were assessed by different testers who likely had different personal styles of interacting with children. Similarly, twins were assessed on different days and events that occurred prior to coming to the lab could impact children's behaviors in the lab and these differences could contribute to nonshared environmental variance. It is important to note that these lab-based explanations cannot account for the substantial nonshared environmental variances for parent-ratings of IC. Differential experiences such as accidents or illnesses can also contribute to nonshared environmental variance.

The bivariate genetic analysis of observer- and parent-assessed IC showed that the two measures tap roughly 47% of the same genetic influences, and that the remaining genetic factors are unique to each measure. Genetic covariance could occur because to some extent the two measures tap the same genetically influenced behaviors (e.g., simple delay ability). The fact that the genetic correlation between observed and parent-rated IC is only moderate may be due to conceptual differences between the two measures (e.g., the TBAQ includes a wider range of behaviors across multiple contexts, and different genetic factors may influence these varied aspects of IC). Although the genetic correlation between the two methods of assessing IC was only moderate, it is this common genetic

covariance that fully accounts for the phenotypic correlation between the two measures of IC. This is typically the case when two variables are significantly heritable but only modestly correlated (Saudino and Plomin 2007).

The finding of genetic variance for IC at 24 months of age is in accordance with temperament theories that presume a biological basis for early appearing individual differences in child temperament (Goldsmith et al. 1987). IC can be viewed as the result of both genes and environments, and both will be taken into account when attempting to modify IC behavior. It is important to note that heritable traits such as IC can be modified by environmental factors. Evidence of genetic, shared (parent-rated IC only) and nonshared environmental influences indicate that behavioral programs and other intervention strategies may be successful in changing maladaptive patterns of IC. However, further research on this phenotype is necessary to clarify the specific mechanisms that may be useful for the modification of IC and related behaviors.

A limitation to the current research is the sample size. Because it is relatively modest in size, we were unable to test for sex differences at the level of etiology. Despite this drawback, the present study includes intensive, laboratory-based measures of IC that are rarely used in studies with larger samples that tend to rely only on parent-rating questionnaires. Behavioral genetic studies of temperament and personality have been criticized for relying on a single measure, usually parent ratings in childhood and self-report in adulthood (Goldsmith et al. 2000; McClelland 1996). The use of multiple sources of information about participants' behavior in a quantitative genetic analysis allows for firmer conclusions about genetic and environmental contributions to the behavior being investigated (Saudino 2005b). Therefore, our consistent findings across two different methods of assessment should be viewed as providing strong initial support for genetic variance on IC. We would caution, however, that although there are some shared genetic influences on both measures, they are to some extent assessing different phenotypes and should not be viewed as fully interchangeable. This is true for both researchers interested in genetic influences and developmentalists more generally.

This research suggests a number of areas for future investigations to pursue. To date, only three studies have explored individual differences in IC. More research is needed with particular attention to possible age differences in the magnitude of genetic and environmental effects and questions about genetic and environmental contributions to age-related continuity and change. Based on previous findings, we know that mean levels of IC increase from infancy to preschool age (Kochanska et al. 1997; Reed et al. 1984; Rothbart 1989b), and then decrease later in adulthood although rank order is stable (Williams et al. 1999). Longitudinal studies of IC can contribute to our knowledge of continuity and change across age, and developmental behavior genetic analyses allow for the estimation of genetic and environmental contributions to stability within the variable. Genetic and environmental influences on specific traits may change across development. Longitudinal developmental behavioral genetic investigations of IC are needed to address this empirical question. Sex differences in genetic and environmental effects are another area for future research. Males and females differed on mean levels of IC in our study. However, the factors that influence mean levels in the population do not necessarily influence individual differences. This is an empirical question that needs large genetically-sensitive samples to answer. Finally, the genetic influences in the present study are anonymous. That is, although quantitative genetic methods such as the twin study can detect genetic variance, these methods do not inform about the specific genes that influence IC. Molecular genetic techniques such as linkage and association studies will be an important next step in the study of IC.

This investigation is a first step in examining genetic and environmental influences on this important behavioral dimension, and findings in this early childhood sample expand on the results of previous studies by providing novel evidence of the specific mechanisms involved in individual differences in IC. The finding that genetic and environmental factors contribute to variability in IC provides important support for current theories of temperament, and bolsters the results of previous investigations that focused on broader temperament factors (e.g., EC), employed only parent ratings, or focused on more complex later developing IC behaviors.

Acknowledgments This research was supported by a predoctoral National Research Service Award (National Institute of Mental Health #1 F31 MH076353-01) and an Elizabeth Munsterberg Koppitz Child Psychology Graduate Fellowship from the American Psychological Foundation (American Psychological Association) awarded to the first author. The Boston University Twin Project (BUTP) is supported by grant MH062375 from the National Institute of Mental Health. Special thanks to the students, staff, parents, and twins from the BUTP.

References

Castellanos FX, Sonuga-Burke EJS, Milham MP, Tannock R (2006) Characterizing cognition in ADHD: beyond executive dysfunction. Trends Cogn Sci 10:117–123

Chiappe P, Hasher L, Siegel LS (2000) Working memory, IC, and reading disability. Mem Cogn 28:8–17

Conway MA, Harries K, Noyes J, Racsma'ny M, Frankish CR (2000) The disruption and dissolution of directed forgetting: IC of memory. J Mem Lang 43:409–430

Deater-Deckard K, Petrill SA, Thompson LA (2007) Anger/frustration, task persistence, and conduct problems in childhood: a behavior genetic analysis. J Child Psychol Psychiatry 48:80–87

Derryberry D, Rothbart MK (1997) Reactive and effortful processes in the organization of temperament. Dev Psychopathol 9:633–652

Eisenberg N, Cumberland A, Spinrad TL, Fabes RA, Shepard SA, Reiser M, Valiente C, Murphy SH, Losoya SH, Guthrie IK (2001) The relations of regulation and emotionality to children's externalizing and internalizing problem behavior. Child Dev 72:1112–1134

Eisenberg N, Spinrad TL, Fabes RA, Reiser M, Cumberland A, Shepard SA, Valiente C, Losoya SH, Guthrie IK, Thompson M (2004) The relations of effortful control and impulsivity to children's resiliency and adjustment. Child Dev 75:25–46

Frye D, Zelaso PD, Palfai T (1995) Theory of mind and rule-based reasoning. Cogn Dev 10:483–527

Garon N, Bryson SE, Smith IM (2008) Executive function in preschoolers: a review using an integrative framework. Psychol Bull 134:31–60

Goldsmith HH (1996) Studying temperament via construction of the Toddler behavior assessment questionnaire. Child Dev 67:218–235

Goldsmith HH, Buss AH, Plomin R, Rothbart MK, Thomas A, Chess S, Hinde RA, McCall RB (1987) Roundtable: what is temperament? Four Approaches. Child Dev 58:505–529

Goldsmith HH, Rieser-Danner LA, Briggs S (1991) Evaluating convergent and discriminant validity of temperament questionnaires for preschoolers, toddlers, and infants. Dev Psychol 27:566–579

Goldsmith HH, Reilly J, Lemery KS, Longley S, Prescott A (1995) The laboratory temperament assessment battery-preschool version: description of procedures. Unpublished manuscript

Goldsmith HH, Buss KA, Lemery KS (1997) Toddler and childhood temperament: expanded content, stronger genetic evidence, new evidence for the importance of environment. Dev Psychol 33:891–905

Goldsmith HH, Lemery KS, Buss KA, Campos J (1999) Genetic analyses of focal aspects of infant temperament. Dev Psychol 35:972–985

Goldsmith HH, Lemery K, Aksan N, Buss KA (2000) Temperamental substrates of personality development. In: Molfese D, Molfese V (eds) Temperament and personality development across the lifespan. Erlbaum, Mahwah, pp 1–32

Griffin D, Gonzalez R (1995) Correlational analysis of dyad-level data in the exchangeable case. Psychol Bull 118:430–439

Groot AS, de Sonneville LMJ, Stins JF, Boomsma DI (2004) Familial influences on sustained attention and inhibition in preschoolers. J Child Psychol Psychiatry 45:306–314

Hughes C, White A, Sharpen J, Dunn J (2000) Antisocial, angry, and unsympathetic: "Hard-to-manage" preschoolers' peer problems and possible cognitive influences. J Child Psychol Psychiatry 41:169–179

Kochanska G, Murray K, Jacques TY, Koenig AL, Vandegeest KA (1996) IC in young children and its role in emerging internalization. Child Dev 67:490–507

Kochanska G, Murray K, Coy KC (1997) IC as a contributor to conscience in childhood: from toddler to early school age. Child Dev 68:263–277

Kochanska G, Murray KT, Harlan ET (2000) Effortful control in early childhood: continuity and change, antecedents, and implications for social development. Dev Psychol 36:220–232

Lemery-Chalfant K, Doelger L, Goldsmith HH (2008) Genetic relations between effortful and attentional control and symptoms of psychopathology in middle childhood. Infant Child Dev 17:365–385

Lengua LJ (2003) Associations among emotionality, self-regulation, adjustment problems, and positive adjustment in middle childhood. Appl Dev Psychol 24:595–618

Liang KY, Zeger SL (1986) Longitudinal data analysis using generalized linear models. Biometrika 73:13–22

Mangelsdorf SC, Schoppe SJ, Buur H (2000) The meaning of parental reports: a contextual approach to the study of temperament and behavior problems in childhood. In: Molfese VJ, Molfese DL (eds) Temperament and personality development across the life span. Lawrence Erlbaum Associates, Mahwah, pp 121–140

McClelland DC (1996) Does the field of psychology have a future? J Res Pers 30:429–434

McGue M, Bouchard TJ Jr (1984) Adjustment of twin data for the effects of age and sex. Behav Genet 14:325–343

Neale MC, Boker SM, Xie G, Maes HH (2003) Mx: statistical modeling, 6th edn. VCU, Department of Psychiatry, Richmond, VA

Neale MC, Stevenson J (1989) Rater bias in the EASI temperament scales: a twin study. J Pers Soc Psychol 56:446–455

Nigg JT, Quamma JP, Greenberg MT, Kusche CA (1999) A two-year longitudinal study of neuropsychological and cognitive performance in relation to behavioral problems and competencies in elementary school children. J Abnorm Child Psychol 27:51–63

O'Connor BP (2004) SPSS and SAS programs for addressing interdependence and basic levels-of-analysis issues in psychological data. Behav Res Methods Instrum Comput 36:17–28

Olson SL, Bates JE, Bayles K (1990) Early antecedents of childhood impulsivity: the role of parent-child interaction, cognitive competence, and temperament. J Abnorm Child Psychol 18:317–334

Olson SL, Schilling EM, Bates JE (1999) Measurement of impulsivity: construct coherence, longitudinal stability, and relationship with externalizing problems in middle childhood and adolescence. J Abnorm Child Psychol 27:151–165

Oosterlaan J, Sergeant J (1996) Inhibition in ADHD, aggressive, and anxious children: a biologically based model of child psychopathology. J Abnorm Child Psychol 24:19–36

Oosterlaan J, Logan GD, Sergeant JA (1998) Response CD, anxious, and control+inhibition in AD/HD, CD, comorbid AD/HD children: a meta-analysis of studies with the stop task. J Child Psychol Psychiatry 39:411–425

Pallodino P, Mammarella N, Vecchi T (2003) Modality-specific effects in inhibitory mechanisms: the interaction of peripheral and central components in working memory. Brain Cogn 53:263–267

Passolunghi MC, Siegel LS (2001) Short-term memory, working memory, and IC in children with difficulties in arithmetic problem solving. J Exp Child Psychol 80:44–57

Perner J, Leekam SR, Wimmer H (1987) Three-year-olds' difficulty with false belief: the case of conceptual deficit. Br J Dev Psychol 5:125–137

Perner J, Stummer S, Lang B (1999) Executive functions and theory of mind: cognitive functioning or functional dependence? In: Zelaso PD, Astington JW, Olson DR (eds) Developing theories of intention: social understanding and self control. Erlbaum, Mahwah, pp 133–152

Pliszka SR, Borcherding SH, Spratley K, Leon S, Irick S (1997) Measuring IC in children. Dev Behav Pediatr 18:254–259

Plomin R, Emde R, Braungart JM, Campos J, Corley R, Fulker DW, Kagan J, Reznick S, Robinson J, Zahn-Waxler C, DeFries JC (1993) Genetic change and continuity from 14 to 20 months: the MacArthur longitudinal twin study. Child Dev 64:1354–1376

Polderman TJC, de Geus EJC, Hoekstra RA, Bartels M, van Leeuwen M, Verhulst FC, Posthuma D, Boomsma DI (2009) Attention problems, inhibitory control, and intelligence index overlapping genetic factors: A study in 9-, 12-, and 18-year-old twins. Neuropsychology 23:381–391

Reed MA, Pien DL, Rothbart MK (1984) Inhibitory self-control in preschool children. Merrill-Palmer Q 30:131–147

Rothbart MK (1989a) Temperament and development. In: Kohnstamm GA, Bates JA, Rothbart MK (eds) Temperament in childhood. Wiley, New York, pp 187–247

Rothbart MK (1989b) Temperament in childhood: a framework. In: Kohnstamm GA, Bates JA, Rothbart MK (eds) Temperament in childhood. Wiley, New York, pp 187–247

Rothbart MK, Ahadi SA (1994) Temperament and the development of personality. J Abnorm Psychol 103:55–66

Rothbart MK, Bates JE (2006) Temperament. In: Eisenberg N (ed) Handbook of child psychology, 6th ed. (Vol 3. Social, emotional, and personality development, pp 99–166). Hoboken, NJ: Wiley

Russell J, Mauthner N, Sharpe S, Tidswell T (1991) The 'windows tasks' as a measure of strategic deception in preschoolers and autistic subjects. Br J Dev Psychol 9:331–349

Saudino KJ (2003) Parent ratings of infant temperament lessons from twin studies. Infant Behav Dev 26:100–107

Saudino KJ (2005a) Rater Bias. In: Everitt B, Howell D (eds) Encyclopedia of statistics in behavioral science, vol 4. Wiley, Chichester, pp 1716–1717

Saudino KJ (2005b) Multiple informants. In: Everitt B, Howell D (eds) Encyclopedia of statistics in behavioral science, vol 4. Wiley, Chichester, pp 1332–1333

Saudino KJ (2009) Do different measures tap the same genetic influences? A multi-method study of activity level in young twins. Dev Sci 12:626–633

Saudino KJ, Cherny SS (2001) Parental ratings of temperament in twins. In: Emde RN, Hewitt JK (eds) Infancy and early childhood: genetic and environmental influences on developmental change. Oxford University Press, New York, pp 73–88

Saudino KJ, Plomin R (2007) Why are hyperactivity and academic achievement related? Child Dev 78:972–986

Saudino KJ, Wertz AE, Gagne JR, Chawla S (2004) Night and day: are siblings as different in temperament as parents say they are? J Pers Soc Psychol 87:698–706

Schachar R, Logan GD (1990) Impulsivity and inhibitory control in normal development and childhood psychopathology. Dev Psychol 26:710–720

Schachar R, Tannock R, Marriott M, Logan G (1995) Deficient IC in attention deficit hyperactivity disorder. J Abnorm Child Psychol 23:411–437

Schachar R, Mota VL, Logan GD, Tannock R, Klim P (2000) Confirmation of an inhibitory control deficit in attention-deficit/hyperactivity disorder. J Abnorm Child Psychol 28:227–235

Scheres A, Oosterlaan J, Sergeant JA (2001) Response execution and inhibition in children with AD/HD and other disruptive disorders: the role of behavioural activation. J Child Psychol Psychiatry 42:347–357

Seifer R, Sameroff AJ, Barrett LC, Krafchuck E (1994) Infant temperament measured by multiple observations and mother report. Child Dev 65:1478–1490

Van Mourik R, Oosterlaan J, Sergeant JA (2005) The Stroop revisited: a meta-analysis of interference control in AD/HD. J Child Psychol Psychiatry 46:150–165

Williams BR, Ponesse JS, Schachar RJ, Logan GD, Tannock R (1999) Development of IC across the life span. Dev Psychol 35:205–213

Wolfe CD, Bell MA (2003) Working memory and IC in early childhood: contributions from physiology, temperament, and language. Dev Psychobiol 44:68–83

Yamagata S, Takahashi Y, Kijima N, Maekawa H, Ono Y, Ando J (2005) Genetic and environmental etiology of effortful control. Twin Res Hum Genet 8:300–306

Zeger SL, Liang KY (1986) Longitudinal data analysis for discrete and continuous outcomes. Biometrics 42:121–130