Development and Psychopathology **29** (2017), 107–120
© Cambridge University Press 2016
doi:10.1017/S0954579415001236

Observed temperament from ages 6 to 36 months predicts parent- and teacher-reported attention-deficit/hyperactivity disorder symptoms in first grade

MICHAEL T. WILLOUGHBY,[a] NISHA C. GOTTFREDSON,[b] CYNTHIA A. STIFTER,[c] AND THE FAMILY LIFE PROJECT INVESTIGATORS
[a]*RTI International;* [b]*University of North Carolina at Chapel Hill; and* [c]*Pennsylvania State University*

Abstract

This study tested the prospective association between observational indicators of temperament, which were obtained across multiple assessments when children were 6–36 months of age, and parent and teacher reports of children's attention-deficit/hyperactivity disorder (ADHD) behaviors, when children were in first grade. Data were drawn from the Family Life Project and included 1,074 children for whom temperament and either parent- or teacher-reported ADHD behavioral data were available. The results of variable-centered regression models indicated that individual differences in temperament regulation, but not temperamental reactivity, was uniquely predictive of parent- and teacher-reported ADHD behaviors. Latent profile analyses were used to characterize configurations of temperamental reactivity and regulation. Person-centered regression models were subsequently estimated in which temperamental profile membership replaced continuous indicators of temperamental reactivity and regulation as predictors. The results of person-centered regression models indicated that temperamental reactivity and regulation both contributed (both alone and in combination) to the prediction of subsequent ADHD behaviors. In general, the predictive associations from early temperament to later ADHD were of modest magnitude (R^2 = .10–.17). Results are discussed with respect to interest in the early identification of children who are at elevated risk for later ADHD.

Attention-deficit/hyperactivity disorder (ADHD) is an early onset, chronic disorder characterized by developmentally inappropriate levels of hyperactive–impulsive and inattentive behaviors that are evident in multiple settings and that result in functional impairments in social, interpersonal, academic, and occupational domains (Fischer, Barkley, Smallish, & Fletcher, 2002, 2007; Klein et al., 2012; Willcutt et al., 2012). Approximately 3%–5% of all school-aged children worldwide meet diagnostic criteria for ADHD (Erskine et al., 2013; Polanczyk, de Lima, Horta, Biederman, & Rohde, 2007). Given the associated health care, educational, parental work loss, and juvenile justice costs, ADHD has been estimated to account for over $42 billion (range = $36–$52 billion in 2005 dollars) in annual societal costs in the United States (Pelham, Foster, & Robb, 2007). Despite the enormous societal costs and individual suffering, efforts directed at the early identification of and intervention/prevention for children at risk for ADHD has been relatively rare, despite the benefits that could accrue from this work (Kupfer et al., 2000; Sonuga-Barke & Halperin, 2010).

Early intervention and prevention of ADHD require the ability to reliably and accurately identify children early in life who are at risk for exhibiting elevated ADHD symptomatology in middle childhood, the time at which ADHD is most often diagnosed. Early risk factors for child ADHD can be grouped into three general categories, including family history of ADHD (unspecified genetic risks), socioeconomic status (e.g., household poverty, household structure, and parental education), and prenatal and perinatal risks (e.g., prenatal stress and dietary factors, drug and toxicant exposures, birth complications, and low birth weight; Das Banerjee, Middleton, & Faraone, 2007; Froehlich et al., 2011; Galera et al., 2011; Sagiv, Epstein, Bellinger, & Korrick, 2013; Thapar, Cooper, Eyre, & Langley, 2013). Limitations of the current knowledge base include the lack of specificity for risks (e.g., low socioeconomic status is not uniquely related to ADHD) and the consideration of risk factors in isolation (few studies have considered family history or candidate genes as modifiers of risks; see, e.g., Nigg, Nikolas, & Burt, 2010). Recent research involving prenatal exposure to smoking as a risk for ADHD serves as a case in point. Whereas prenatal exposure to smoking was previously considered a risk factor for ADHD (Linnet et al., 2003; Thapar et al., 2003), subsequent studies have indicated that the previous associations may have

Support for this research was provided by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (Grants P01-HD039667 and R03-HD71646), with cofunding from the National Institute on Drug Abuse. The Family Life Project Key Investigators in the second phase of the study included Lynne Vernon-Feagans, Martha Cox, Clancy Blair, Peg Burchinal, Linda Burton, Keith Crnic, Patricia Garrett-Peters, Mark Greenberg, Roger Mills-Koonce, and Michael Willoughby.

Address correspondence and reprint requests to: Michael Willoughby, RTI International, Room 349 Hobbs Building, 3040 Cornwallis Road, Research Triangle Park, NC 27709-2194; E-mail: mwilloughby@rti.org.

been spurious (Skoglund, Chen, D'Onofrio, Lichtenstein, & Larsson, 2014; Thapar et al., 2009) or that the risk for ADHD that results from prenatal exposure to smoking may be conditional on specific genetic characteristics of exposed children (Altink et al., 2008; Langley, Heron, Smith, & Thapar, 2012; Neuman et al., 2007). While the current status of knowledge on risk factors may facilitate the overselection of children who at increased risk for ADHD into research studies, it is unlikely to be useful for the early identification of individual children who are at risk for subsequent ADHD. This undermines the ability to develop and test models of early intervention that are intended to either prevent or reduce the severity of subsequent ADHD.

Individual differences in temperament represent another early risk factor that may be useful for identifying children early in life who are at elevated risk for subsequent ADHD. A large literature has considered children's temperament as a vulnerability factor to psychopathology in general (Nigg, 2006). Although early notions of "difficult" temperament (Thomas & Chess, 1977) have had an enduring appeal (Bussing, Lehninger, & Eyberg, 2006; Guerin, Gottfried, & Thomas, 1997; Oberklaid, Sanson, Pedlow, & Prior, 1993), the modern literature has emphasized the joint contributions of reactive (e.g., exuberance, fear, or anger) and regulatory (e.g., effortful control) aspects of temperament to inform psychopathology (Muris & Ollendick, 2005; Stifter & Dollar, in press). With respect to ADHD, Martel, Nigg, and their colleagues proposed a dual-process model in which reactive (i.e., "bottom-up, affective") processes were linked to hyperactive–impulsive symptomatology and regulatory (i.e., "top-down, control") processes were linked to inattentive symptomatology (Martel, Gremillion, & Roberts, 2012; Martel & Nigg, 2006; Martel, Nigg, & Lucas, 2008; Martel, Nigg, & von Eye, 2009). Their model is consistent with other studies that have established that ADHD is associated with temperamental dimensions of overactivity, negative affect, and low persistence/task orientation (Bell, Kellison, Garvan, & Bussing, 2010; Foley, McClowry, & Castellanos, 2008; Healey, Marks, & Halperin, 2011; Lambert, 1982; McIntosh & ColeLove, 1996; Rettew, Copeland, Stranger, & Hudziack, 2004).

Despite the general consistency of results, many previous studies were limited by single informant ratings of highly similar behaviors to represent ADHD and temperament (Lahey, 2004). Moreover, many previous studies focused on *contemporaneous* associations of ADHD and temperament, both of which were first assessed in middle childhood. As such, these studies were uninformative about whether individual differences in temperament in the first years of life were prospectively related to ADHD. However, this is the type of evidence that would be required if temperament were to facilitate the early identification of and intervention for children who at risk for ADHD. Studies that use prospective longitudinal designs and that relate early objective measures of temperament to later ADHD are essential for testing what has increasingly become the "conventional wisdom" regarding the contributions of temperament to ADHD. This is the overarching objective of the current study.

A small number of studies have considered prospective associations between "regulatory" (i.e., persistent crying, sleeping, and/or feeding) problems in the first year of life and subsequent risk for behavioral outcomes, including ADHD. A meta-analysis indicated that early problems with persistent crying (Cohen $d = 0.42$, based on six studies) and especially sleeping (Cohen $d = 1.3$, based on two studies) were prospectively associated with ADHD (Hemmi, Wolke, & Schneider, 2011). In addition, at least one study used a retrospective design to identify infant and toddler characteristics (obtained from medical chart review) that differentiated children with ADHD from age- and gender-matched controls. ADHD children were more likely to have been characterized as having a difficult temperament (restless, irritable, easy frustrated, and/or difficulty postponing immediately satisfaction) by their parents and/or pediatricians when they were 9 and 18 months olds (Gurevitz, Geva, Varon, & Leitner, 2014). Although sleep dysregulation, persistent crying, and general notions of child difficulty do not conform to modern instantiations of temperament (i.e., reactivity and regulation), these studies are important because they establish the plausibility of using objective indicator of children's temperament as a strategy for identifying children who are at risk for school-age ADHD.

ADHD has long been understood to represent a heterogeneous disorder, which raises the prospect that there may be distinct profiles of temperamental risk among ADHD youth. Nigg, Goldsmith, and Sachek (2004) proposed five profiles. While two profiles reflected temperamental main effects of reactivity (1 = *very high approach/activity level*, 2 = *very low regulation*), three reflected temperamental interactions between reactivity and regulation (3 = *very high anger/low regulation*, 4 = *very low fear/high anger/low regulation*, and 5 = *high fear/low anger/low regulation*). All of the proposed profiles identified infancy and toddlerhood as the key developmental periods in which individual differences in temperament would first emerge.

Inspired by Nigg et al. (2004), Karalunas et al. (2014) recently subdivided a large sample of ADHD youth, who ranged from 7 to 11 years of age, into three distinct temperamental subtypes (mild, exuberant, and irritable), each of which exhibited unique physiological, neurological, and clinical correlates. This study provided some of the first empirical support for the value of using a person-oriented approach to represent temperamental heterogeneity within a population of ADHD youth. However, similar to other studies that were reviewed above, this study was limited to the measurement of temperament after the onset of diagnosis.

One of the main objectives of the current study was to test prospective associations between observational measures of temperament obtained in infancy and toddlerhood and the emergence of parent- and teacher-reported ADHD behaviors in children. The use of a large sample, a prospective design, modern characterizations of temperament (reactivity and regulation), and both variable-centered (i.e., the use of temperament scores as continuous indicators) and person-oriented (i.e., latent profiles of combinations of temperamental indicators) approaches for representing temperament scores all represented important extensions and innovations of previous studies. Strong tests of whether early objective indicators of temperament were prospectively association with school-aged ADHD behaviors informs questions about whether temperament could be used to facilitate early identification and intervention for ADHD (see Stifter & Dollar, in press).

ADHD is a familial disorder (Biederman et al., 1992; Epstein et al., 2000). Family history of ADHD is often characterized as a nonspecific marker of genetic risk for child ADHD (Stawicki, Nigg, & von Eye, 2006; Thapar et al., 2013). There is evidence that one of the ways in which a family history of ADHD increases risk for child ADHD is through its impact on the neuropsychological processes, including executive functions and delay aversion (Pauli-Pott, Dalir, Mingebach, Roller, & Becker, 2013; Seidman et al., 1995). Sullivan et al. (2015) also recently reported that a family history of ADHD was also associated with objective measures of children's early temperamental reactivity. These findings build on the work of Auerbach et al. (2005, 2008) and Landau et al. (2010), who demonstrated that a family history of ADHD (i.e., biological father ADHD symptomatology that was assessed when children were born) was associated with differences in multiple dimensions of children's temperament (state organization, activity level, anger, and attentional control) at assessments spanning 7 weeks through 25 months. Family history of ADHD was also predictive of inhibitory control, and the combination of family history and a risk polymorphism of the dopamine receptor D4 gene (i.e., *DRD4*) were predictive of ADHD symptoms at age 4.5 years (Auerbach et al., 2010; Berger, Alyagon, Hadaya, Atzaba-Poria, & Auerbach, 2013). Taken together, these studies raise the possibility that the prospective association between early indicators of temperament and subsequent ADHD behaviors may be stronger for children with a positive family history of ADHD. Given the limited nature of directly supporting evidence, this question was conceived of as more hypothesis generating than hypothesis testing.

ADHD has long been understood to be disproportionally more common among males than among females; however, the magnitude of the discrepancy is less pronounced in community (vs. clinical) samples and may be specific to middle childhood (Ramtekkar, Reiersen, Todorov, & Todd, 2010; Rucklidge, 2010). Although males and females who meet diagnostic criteria for ADHD have been described as similarly impaired in multiple domains of functioning (Biederman & Faraone, 2004; Levy, Hay, Bennett, & McStephen, 2005; Seidman et al., 2005; Yoshimasu et al., 2010), there are a few indications that specific risk factors for ADHD may vary by gender (Froehlich et al., 2007; Hermens et al., 2004; Martel, 2013; Martel, Lucia, Nigg, & Breslau, 2007). Gender differences are also evident in multiple dimensions of temperament that are implicated in ADHD (Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006). Small- to moderate-sized sex differences in observed anger reactivity, activity level, and regulation were recently reported in the same sample that will be used in the current study (Willoughby, Stifter, Gottfredson, & the Family Life Project Investigators, 2014). Similar to family history of ADHD, we tested whether any prospective association between early indicators of temperament and subsequent ADHD behaviors may be stronger for boys than for girls. Once again, given the limited nature of directly relevant research, we considered child gender an exploratory moderator.

In sum, the current study tested whether multiple dimensions of temperament, which were measured across the first 3 years of life, predicted parent and teacher reports of ADHD symptomatology at first grade. Dimensions of temperamental reactivity (fear, anger, and activity level; measured from 6 to 24 months) and regulation (inhibitory control and attention/persistence; measured from 24 to 36 months) were jointly (and multiplicatively) considered as predictors of ADHD using both variable-oriented and person-centered methods. By relying exclusively on observational and direct performance measures of temperament and parent and teacher ratings of ADHD symptoms, we eliminated method effects as a potential confounding variable. Family history of ADHD and child gender were both considered potential exploratory moderators of the association between early temperament and later ADHD. The broader objective of this study was to contribute to ongoing efforts to identify early risk factors for ADHD that would facilitate early identification and intervention efforts.

Methods

Participants and procedures

The Family Life Project is a prospective longitudinal study of families residing in six low-wealth counties in eastern North Carolina and central Pennsylvania (three counties per state) that were selected to be indicative of the "Black South" and northern Appalachia, respectively. Complex sampling procedures were employed to recruit a representative sample of 1,292 children whose families resided in one of the six counties at the time of the child's birth. Low-income families in both states and, in North Carolina, African American families were oversampled; however, through the use of weighted analyses, all of our inferences generalize back to the six-county study area as if participants were selected using simple random sampling. Full details of the sampling plan and study design appear elsewhere (Vernon-Feagans, Cox, & the Family Life Project Investigators, 2013).

The current study made use of temperament factor scores that were derived from data that were collected at the 6-, 15-, 24-, and/or 36-month home visits. Factor scores representing three dimensions of reactivity (activity level, fear, and anger; all based on data from 6 to 24 months) and a dimension of regulation (based on data from 24 and 36 months) were created using indicators that included microsocial (second-by-second coding) and macrosocial (home visitor ratings of children's behavior across the entire home visit) observations, as well as children's performance on direct assessment measures. Temperamental factor scores were available for $N = 1,205$ (93% of total sample) children. The current study also made use of parent ($N = 1,067$) and teacher ($N = 892$) ratings of ADHD symptoms when target children were in first grade. In total, 1,074 children had temperament and ADHD data from at least one informant. Participating families and children ($N = 1,074$) did not differ from nonparticipating families and children ($N = 218$) with respect to being recruited into the low-income stratum (78% vs. 74%, $p = .15$), state of residence (60% vs. 61% resided in North Carolina, $p = .59$), primary caregiver educational status at study enrollment (80% vs. 80% with a high school degree/GED or beyond, $p = .97$), or child gender (50% vs. 53% male, $p = .44$). Racially, there was a trend for study participants to be more likely African American than nonstudy participants (44% vs. 37%, $p = .06$).

Measures

Temperament factor scores. We recently used confirmatory factor analysis to build constructs of temperamental reactivity and regulation that consisted entirely of observational indicators that were not subject to the limitations of parent reports (Willoughby, Stifter, Gottfredson, & the Family Life Project Investigators, 2015). A four-factor model (activity level, anger reactivity, fear reactivity, and regulation) provided a good fit to observational data, which included 22 indicators that were obtained from home visits that occurred when children were 6–36 months old. A complete description of the measures that were used as indicators of the four temperament factors appear in Appendix A. Factor scores for each of the four temperament constructs were used as the primary independent variables in this study. Factor scores were standardized to $M = 0$ and $SD = 1$ to improve interpretability (a one-unit change corresponded to a 1 *SD* increase in a given dimension of temperament).

The activity level factor score was based on four indicators: an observation of activity level during a free-play interaction with a primary caregiver at the 6-month home visit and combined home visitor ratings of gross motor behavior (i.e., "amount of gross bodily movement") that were made at the end of the 6-, 15-, and 24-month home visits. The anger reactivity factor score was based on four indicators: observed negative affect in response to the Laboratory Assessment Battery (Lab-TAB; Goldsmith & Rothbart, 1996) toy removal task and combined home visitor ratings (i.e., "irritability"), both of which were completed at the 15- and 24-month home visits. The fear reactivity factor score was based on four indicators: observed fear in response to the Lab-TAB masks tasks and combined home visitor ratings (i.e., "reaction to new or strange"), both of which were completed at the 15- and 24-month home visits. The regulation factor score was based on 10 indicators: combined home visitor ratings of persistence at the 24-month (i.e., "tendency to persist in attention to any one objective, person, or activity, aside from attaining a goal") and 36-month (i.e., "task persistence–degree of on-task behavior and persistence in the face of frustration") assessments, observed persistence during parent–child interactions involving a joint puzzle-solving task at each of the 24- and 36-month home visits, and child performance on two direct assessments at the 24-month home visit (i.e., delay of gratification and reverse categorization) and four direct assessments at the 36-month home visit (i.e., silly sounds Stroop, go/no-go, spatial conflict, and something the same) that collectively measured inhibitory control and attention shifting.

Child ADHD symptom ratings. Parents and teachers independently completed an ADHD rating scale at the first-grade visit (Pelham, Evans, Gnagy, & Greenslade, 1992). All 18 DSM-IV symptoms for ADHD were rated on a 4-point Likert-like scale (0 = *not at all*, 1 = *just a little*, 2 = *pretty much*, and 3 = *very much*). Consistent with our earlier work that involved this sample, which demonstrated the central importance of overall ADHD relative to inattentive and hyperactive–impulsive scores (Willoughby, Blanton, & the Family Life Project Investigators, 2015; Willoughby, Pek, Greenberg, & the Family Life Project Investigators, 2012), we focused on total ADHD scores in the current study. In order to preserve as much variability in ADHD symptoms as possible (which benefitted prediction), a total mean ADHD score (i.e., the average of all 18 items, each rated on the 0–3 Likert format) was created separately for parents ($\alpha = 0.94$) and teachers ($\alpha = 0.96$). Parent and teacher total ADHD scores served as the two primary outcomes.

Family history of ADHD. A single item was asked to establish whether either the biological mother or father of the target child had a childhood history of ADHD (i.e., "Has a doctor or other medical professional ever told you [him/her] that you [s/he] have [has] attention-deficit disorder"). When the respondent was the biological mother of the target child, which was typically the case, she answered the question about herself and the child's biological father. When the respondent was the biological father of the target child, he answered the question about himself and the child's biological mother. When the respondent was not a biological parent of the target child, s/he answered the question with reference to the child's biological parents. In total, 2.7% (29/1073) of biological mothers, 3.8% (39/1036) of biological fathers, and 5.9% of either biological mothers or fathers (63/1073) were identified as having a positive family history. Data was completely missing for one family; moreover, there was missing data from some biological fathers for whom the mother or respondent could not provide a definitive answer. Given the low base rates of individual history for mothers and fathers, the combined ("either") item was used as the predictor in the current study.

Analytic strategy

This study used variable- and person-centered approaches to test whether four dimensions of temperament (activity level,

https://doi.org/10.1017/S0954579415001236 Published online by Cambridge University Press

fear and anger reactivity, and regulation), which were derived from multiple observational and direct performance tasks when children were 6–36 months old, were predictive of parent and teacher reports of ADHD when children were in first grade. Variable-centered models regressed parent and teacher reports of ADHD (in separate models) onto the four dimensions of temperament, child gender, and family history of ADHD. Interactions between reactivity and regulation (i.e., three terms), gender and temperament (i.e., four terms), and family history of ADHD and temperament (i.e., four terms) were also considered to test whether specific dimensions of temperament were differentially predictive as a function of gender or family history. Interaction terms were omitted from the final model if they were not statistically significant. Gender (male) and family history (positive) were mean centered. Factor score estimates of temperament were standardized prior to modeling, while mean levels of ADHD behaviors were left in their original scale. Hence, the reported regression coefficients represented the expected mean shift in parent- and teacher-reported mean ADHD scores (rated 0–3 metric) for each 1 *SD* shift in temperamental factor scores. This approach also facilitated the computation of effect sizes, which are reported in the text.

Person-oriented models initially involved estimating a series of latent profile models in which the four dimensions of temperament (i.e., factor scores) were the indicators. Following best practice (Nylund, Asparouhov, & Muthen, 2008; Tofighi & Enders, 2007), a combination of statistical (i.e., Akaike, Bayesian, and sample-size adjusted Bayesian information criteria; Lo–Mendal–Rubin test statistic) and pragmatic (e.g., size of the group and theoretical interpretability of temperamental scores within each group) criteria were used to determine the optimal number of groups to represent children's temperament. Regression models were subsequently reestimated, replacing continuous indicators of temperament (including Reactivity × Regulation interactions) with dummy-coded variables that represented latent profile membership (where children were assigned to that latent profile group for which they had the largest posterior probability). The group that exhibited the most normative temperament profile was designated the reference group. Analogous to the variable-centered models, interactions between latent profile group membership and gender and family history were separately considered to test whether specific profiles of temperament were differentially predictive of ADHD. Interaction terms were omitted from the final model if they were not statistically significant.

All regression models were estimated using PROC SURVEYREG in SAS® version 9.3 to accommodate the complex sampling design (i.e., stratification and oversampling for low-income and, in North Carolina, African American families). Given the large number of interactions tested, we set an a priori threshold of $p < .01$ in order to control for Type I errors. Multiple imputation procedures (as implemented using PROCs MI and MIANALYZE in SAS version 9.3) were used to address missing data in all regression models, which primarily consisted of missing teacher-rated ADHD data.

Results

Descriptive statistics

Descriptive statistics, including bivariate correlations, for all study variables are summarized in Table 1. Four points were noteworthy. First, bivariate correlations among the temperament variables were of moderate magnitude ($|rs| = .08–.57$). Second, anger reactivity and regulation, but not activity and fear reactivity, were moderately associated with parent- and teacher-reported ADHD behaviors in first grade ($|rs| = .21–.35$). Third, family history had a small but significant positive association with both parent-reported and teacher-reported ADHD behaviors. Fourth, boys were reported as having higher mean levels of ADHD behaviors, especially per teacher reports.

Variable-centered analyses

A series of regression models was estimated for parent and teacher ratings of ADHD behaviors. The initial model in-

Table 1. *Unweighted descriptive statistics for predictor and outcome variables*

	1	2	3	4	5	6	7	8
1. Activity level (FS)	—							
2. Fear reactivity (FS)	−.08*	—						
3. Anger reactivity (FS)	.24***	.52***	—					
4. Regulation (FS)	.18***	−.18***	−.57***	—				
5. Male	.10***	.03	.17***	−.21***	—			
6. Family history ADHD	−.06	−.05	.01	−.09**	.00	—		
7. Parent reported ADHD	−.03	.07*	.21***	−.35***	.10***	.13***	—	
8. Teacher reported ADHD	.00	.13***	.23***	−.33***	.25***	.07*	.39***	—
N	1,074	1,074	1,074	1,074	1,074	1,073	1,067	892
M	0	0	0	0	0.50	0.06	0.82	0.71
SD	1	1	1	1	0.50	0.24	0.62	0.76

Note: FS, Factor score; ADHD, attention-deficit/hyperactivity disorder. Parent- and teacher-rated ADHD mean scores range from 0 to 3.
*$p <. 05$. **$p < .01$. ***$p < .001$.

https://doi.org/10.1017/S0954579415001236 Published online by Cambridge University Press

cluded six main effects (four dimensions of temperament: activity level, fear, anger, and regulation; as well as child gender and family history of ADHD). As described above, subsequent models investigated three sets of interaction terms including Regulation × Reactivity (i.e., three interaction terms: Activity × Regulation, Fear × Regulation, and Anger × Regulation), Gender × Temperament (i.e., four interaction terms: Activity × Male, Fear × Male, Anger × Male, and Regulation × Male), and Family History of ADHD × Temperament (i.e., four interaction terms: Activity × Family History, Fear × Family History, Anger × Family History, and Regulation × Family History), with each set being considered separately. Because none of these interactions was statistically significant after omission of the other nonsignificant terms, these results were not presented below.

Parent-rated ADHD. Children with higher levels of regulation were reported to exhibit lower mean levels of ADHD behaviors. Specifically, a 1 *SD* increase in regulation was associated with a 0.24 point decrease in parent-rated ADHD behaviors ($b = -0.24$, $p < .0001$). None of the three dimensions of reactivity were uniquely related to parent-rated ADHD behaviors (see Table 2). In addition, children whose parent had a childhood history of ADHD were also rated by their parent as having higher levels of ADHD behaviors ($b = 0.26$, $p = .002$). Given that the standard deviation of parent-rated ADHD was 0.62 (Table 1), temperamental regulation and family history represented moderate-sized effects (Cohen *ds* ≈ 0.40). Collectively, temperamental, family history, and child gender predicted 17% of the variation in parent-rated ADHD behaviors ($R^2 = .17$).

Teacher-rated ADHD. Children with higher levels of regulation were reported to exhibit lower mean levels of ADHD behaviors. Specifically, a 1 *SD* increase in regulation was associated with a 0.24 point decrease in teacher-rated ADHD behaviors ($b = -0.24$, $p < .0001$). None of the other three dimensions of reactivity were uniquely related to teacher-rated ADHD behaviors (see Table 2). Although family history was not predictive of teacher-rated ADHD ($b = 0.16$, $p = .13$), child gender was ($b = 0.24$, $p < .0001$). Boys exhibited higher mean levels of teacher-rated ADHD behaviors than girls. Given that the standard deviation of teacher-rated ADHD was 0.76 (Table 1), temperamental regulation and gender represented small- to moderate-sized effects (Cohen *ds* ≈ 0.33). Collectively, temperamental, family history, and child gender predicted 15% of the variation in teacher-rated ADHD behaviors ($R^2 = .15$).

Table 2. *Summary of regression coefficients for parent- and teacher-rated ADHD behaviors in first grade for the variable-centered approach*

	Parent Report	Teacher Report
	b (*SE*)	*b* (*SE*)
Intercept	0.82*** (0.02)	0.67*** (0.03)
Male	0.03 (0.04)	0.23*** (0.05)
Regulation	−0.25*** (0.03)	−0.24*** (0.03)
Anger reactivity	−0.02 (0.03)	−0.05 (0.04)
Fear reactivity	0.00 (0.02)	0.06 (0.03)
Activity	0.03 (0.02)	0.04 (0.03)
Family history ADHD	0.26** (0.08)	0.16 (0.10)
R^2	.17	.15

Note: $N = 1{,}074$. The temperament scores were standardized prior to estimation; corresponding regression coefficients represent the expected change in ADHD symptom severity (rated on a 0–3 scale) given a 1 *SD* increase in temperament. R^2, The average proportion of variance explained across 40 imputation models; ADHD, attention-deficit/hyperactivity disorder.
$^{**}p < .01$. $^{***}p < .001$.

Table 3. *Summary of regression coefficients for parent- and teacher-rated ADHD behaviors in first grade for the person-centered approach*

	Parent Report	Teacher Report
	b (*SE*)	*b* (*SE*)
Intercept	0.66*** (0.03)	0.47*** (0.04)
Male	0.12** (0.04)	0.30*** (0.05)
Family history ADHD	0.32*** (0.09)	0.19 (0.11)
Low anger, low activity	0.09 (0.08)	0.15 (0.12)
High fear, high anger, low regulation	0.36*** (0.06)	0.43*** (0.08)
Low fear, low regulation	0.32*** (0.08)	0.45*** (0.10)
Moderate anger, moderate activity	0.22*** (0.07)	0.29*** (0.08)
Moderate fear, moderate regulation	0.04 (0.05)	0.09 (0.06)
R^2	.09	.12

Note: $N = 1{,}074$. Temperamental profile groups were dummy coded such that they represent differences in rated ADHD behaviors relative to the reference group (i.e., the combined low fear and anger reactivity and high regulation classes which constituted 34% of the population). R^2, The average proportion of variance explained across 40 imputation models; ADHD, attention-deficit/hyperactivity disorder.
$^{**}p < .01$. $^{***}p < .001$.

Person-centered analyses

In order to facilitate person-centered analyses, a series of latent profile models were estimated that used temperament factor scores as indicators. All of the information criteria indicated that each additional class improved model fit (e.g., Bayesian information criterion values for 1–8 classes decreased monotonically: 12,184, 11,409, 11,088, 10,962, 10,908, 10,830, 10,794, and 10,777). By the time eight classes were extracted, class sizes became too small (1% of the population) to be useful. A close inspection of the three- to eight-class solutions indicated a high degree of similarity across models. The seven-class solution was chosen because it represented the greatest degree of variation in temperamental profiles while retaining sample sizes that were large enough to be useful (>5% of the sample) and because the estimated classes mapped well to our theoretical expectations.

https://doi.org/10.1017/S0954579415001236 Published online by Cambridge University Press

Children were assigned to profiles based on posterior probabilities (i.e., each child was assigned to that class that was most likely to have given risk to their observed data). Figure 1 provides the mean factors scores for each dimension of temperament separately for each of the seven latent profile classes (note that factor scores are interpretable on a z score metric; hence, values of 0 in Figure 1 represent weighted sample average behaviors). Two classes, which were similarly characterized by low fear and anger reactivity and high regulation and which together constituted approximately one third of the sample, were combined to form a reference group. These two groups were combined because they differed on in level not type of temperament (i.e., quantitative not qualitative differences) and because this shared profile of temperament was not considered a risk for ADHD (collapsing these groups did not exactly replicate group assignment from the six-class model). The remaining two-thirds of children were assigned to the five remaining profiles groups. These profile groups differed appreciably in terms of their combinations of reactivity and regulation. Next, the parent and teacher regression models were reestimated such that children's membership in latent profile classes replaced the continuous indicators of temperament (including Reactive × Regulation interactions) as the primary predictors, using the normative group as the reference group.

Similar to the variable-centered regression models, we tested whether the effects of temperamental profile membership in the prediction of parent- and teacher-rated ADHD behaviors differed as a function of child gender or family history of ADHD. This was accomplished by estimating models that included product interaction terms between each temperamental profile group (less the reference group) and family history or child gender (in separate models). Because none of these interactions was statistically significant, these results were not presented below.

Parent-rated ADHD. Relative to children in the reference group, children in the high fear/high anger/low regulation ($b = 0.36$, $p < .0001$), low fear/low regulation ($b = 0.32$, $p < .0001$), and moderate anger/moderate activity ($b = 0.22$, $p = .0008$) classes were all reported to exhibit elevated levels of parent-reported ADHD behaviors. In contrast, children assigned to the low anger/low activity ($b = 0.09$, $p = .31$) and the moderate fear/moderate regulation ($b = 0.04$, $p = .42$) did not. Boys ($b = 0.12$, $p = .004$) and children with a family history of ADHD ($b = 0.32$, $p = .0006$) also exhibited elevated levels of parent-rated ADHD behaviors. Given that the standard deviation of parent-rated ADHD was 0.62 (Table 1), these represented small- to moderate-sized effects (Cohen ds ≈ 0.20–060). Collectively, latent profile membership, family history, and child gender explained 9% of the variation in parent-rated ADHD behaviors ($R^2 = .09$), which was less than variable-centered models.

Teacher-rated ADHD. Relative to children in the reference group, children in the high fear/high anger/low regulation ($b = 0.43$, $p < .0001$), low fear/low regulation ($b = 0.45$, $p < .0001$), and moderate anger/moderate activity ($b = 0.29$, $p = .0006$) classes were all reported to exhibit elevated levels of teacher-reported ADHD behaviors. In contrast, children assigned to the low anger/low activity ($b = 0.15$, $p = .20$) and the moderate fear/moderate regulation ($b = 0.09$, $p = .18$) did not. Boys ($b = 0.30$, $p < .0001$), but not children with a family history of ADHD ($b = 0.19$, $p = .09$), also exhibited elevated levels of teacher-rated ADHD behaviors. Given that the standard deviation of teacher-rated ADHD was 0.76 (Table 1), these represented moderate-sized effects (Cohen ds ≈ 0.40–0.60). Collectively, latent profile membership, family history, and demographic factors explained 11% of the variation in teacher-rated ADHD behaviors ($R^2 = .12$), which was again less than variable-centered models.

In an effort to place these person-centered results into better clinical context, we plotted the number of combined parent- and teacher-reported symptoms (where a symptom was considered present if it was endorsed as "pretty much" or "very much" by either informant) for each temperamental profile group. We also plotted the proportion of children in each temperamental profile group who exhibited six or more inattentive or six and/or more hyperactive–impulsive symptoms, which was intended to approximate diagnostic eligibility (these are upper-bound estimates because they do not take into account impairment in functioning or other diagnostic rule outs). Consistent with the person-centered regression results, specific temperamental profile groups exhibited elevated levels of combined informant ADHD symptoms and diagnostic risk (see Figure 2).

Discussion

Conventional wisdom holds that ADHD is a lifelong disorder that has early manifestations of high levels of temperamental reactivity and low levels of temperamental regulation. However, much of this wisdom is derived from anecdotal reports and cross-sectional studies that considered contemporaneous associations between temperamental and ADHD behaviors. Relatively few studies have tested prospective longitudinal associations between objective indicators of temperament that were collected during the infancy and toddler period and children's ADHD symptomatology at school age. Even fewer studies have been in a position to use modern conceptualizations of temperament as individual differences in reactivity and regulation. From this broader context, the current study made multiple contributions to the literature. First, we established that multiple indicators of temperamental reactivity and regulation, which were measured across the first 3 years of children's life, were associated with parent- and teacher-rated ADHD symptom behaviors when children were in first grade. However, the specific nature of these predictive associations depended on the analytic approach that was used. Second, we provided empirical support for theoretical speculations regarding the heterogeneity of temperamental profiles that are associated with later ADHD. Third, in

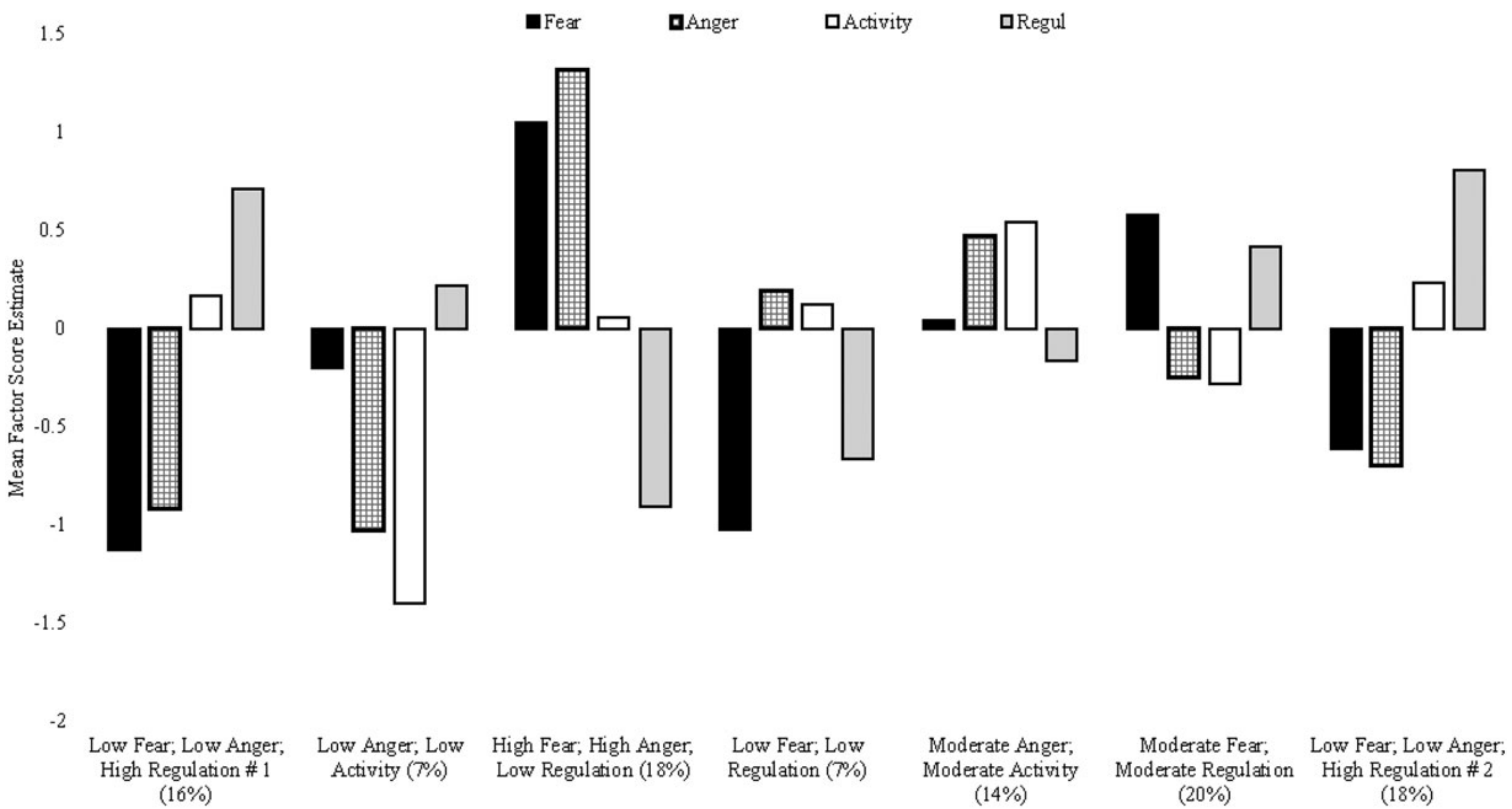

Figure 1. Observed temperament factor scores for temperamental profile groups.

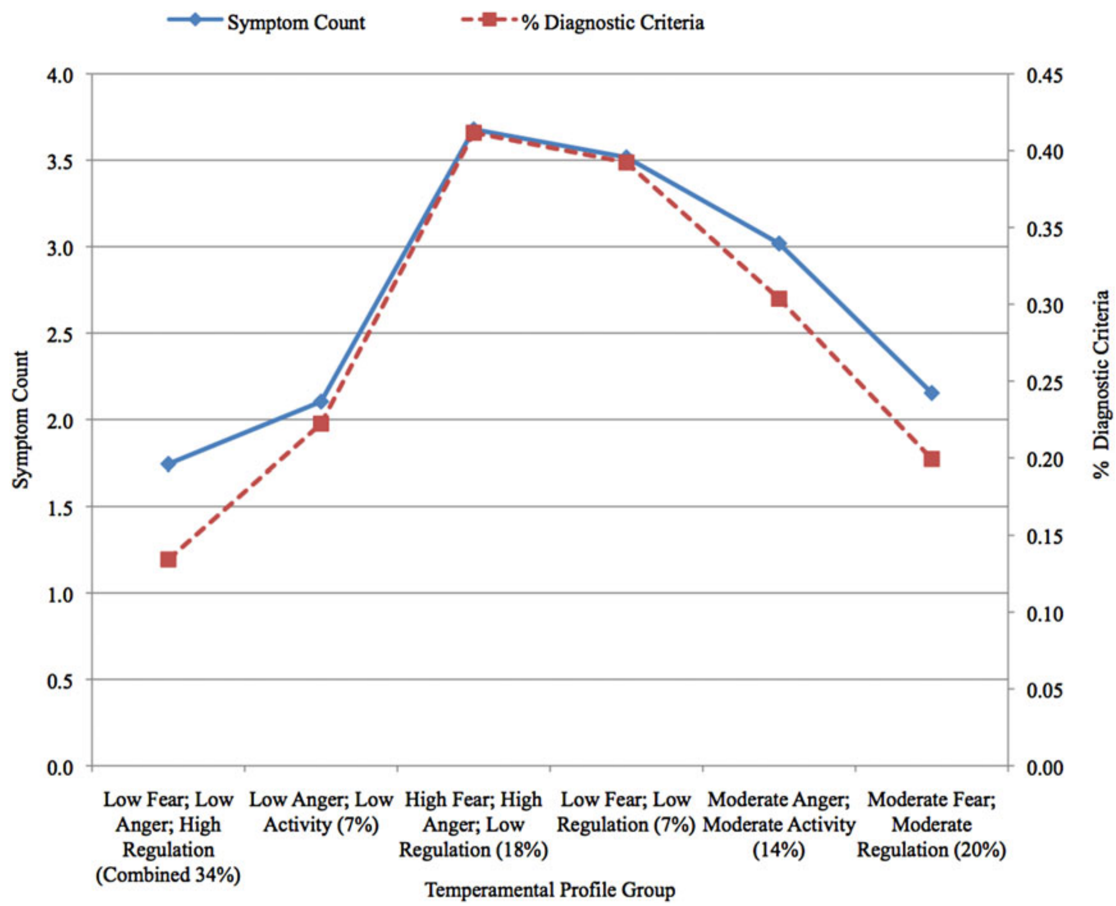

Figure 2. (Color online) Combined parent- and teacher-reported attention-decifit/hyperactivity disorder for temperament profile groups.

contrast to previous smaller scale studies that reported differential effects of temperament on ADHD as a function of family history of ADHD (unspecific genetic risk), we did not find any evidence for moderation. Fourth, although we established prospective associations from regulation in early childhood to ADHD behaviors in first grade, the magnitude of effects was modest. Each of these findings is addressed in turn.

The variable-centered models provided clear support for the unique and central contribution of temperamental regulation in the prediction of ADHD. The variable-centered models did not provide any evidence to suggest that the effects of regulation on later ADHD behaviors were conditional on any dimension of reactivity (i.e., none of the Reactivity × Regulation interaction terms was statistically significant). These results conform to long-standing theory and a large body of evidence that links individual differences in early regulatory capacity to subsequent risk for ADHD problems.

However, a different set of conclusions emerged from the person-centered models. Over a decade ago, Nigg et al. (2004) speculated that there were five temperamental profiles that were associated with increased risk for ADHD. Two profiles were characterized by temperamental main effects (i.e., 1 = *very high approach/activity level* and 2 = *very low regulation*), while three profiles were characterized by temperamental interactions involving reactivity and regulation (3 = *very high anger/low regulation*, 4 = *very low fear/high anger/low regulation*, and 5 = *high fear/low anger/low regulation*). The results of our latent profile analyses conform rather closely to the profile types that were proposed by Nigg et al. (2004). For example, we identified one of Nigg et al.'s main effect groups (14% of the population was characterized by elevated activity level and anger) and two of their interaction groups (18% of the population was characterized by high fear/high anger/low regulation and an additional 7% was characterized by low fear/low regulation). All three of these groups exhibited elevations in both parent- and teacher-rated ADHD behaviors, relative to children in a normative profile that was characterized by low reactivity and high regulation. These person-centered results provided an important elaboration of the variable-centered results. The variable-centered results strongly suggested that after taking into account temperamental regulation, temperamental reactivity was not relevant to later ADHD. Although the person-centered results also revealed the importance of low regulation and increased risk for ADHD, they demonstrated that the low-regulation groups differed in terms of their pattern of temperamental reactivity (low fear/low regulation vs. high fear/high anger/low regulation). However, we did not find a profile of children who were characterized solely by low regulation. Moreover, the person-centered results also shed light

https://doi.org/10.1017/S0954579415001236 Published online by Cambridge University Press

on the risk for ADHD that was evident from the 14% of children who were distinguished solely on the basis of their temperamental reactivity (moderate anger/moderate activity).

Although a positive family history of ADHD was a significant predictor of parent-, but not teacher-, reported ADHD behaviors, we did not find any evidence that a family history of ADHD modified the predictive association between early temperament and later ADHD behaviors. Two points are noteworthy. First, we designated children as having a positive family history of ADHD on the basis of either their biological mother's or father's history. This decision was pragmatic, because the base rates of family history for biological mothers and fathers were low. Initially, we considered the separate effects of mother and father history, with some evidence that a father history of ADHD may interact with temperament to predict ADHD behaviors. However, these results did not withstand sensitivity checks (i.e., results were unstable), which we attributed to small sample sizes. This suggests that studies that use family history of ADHD as a proxy to general genetic risk would be well served by using a stratified sampling plan in order to enrich the sample for adults with a history of ADHD. Second, our reliance on family history may be too general. Becker and colleagues demonstrated that the predictive association between early regulatory problems (e.g., crying and sleeping) and later ADHD was specific to children with a seven-repeat allele of the dopamine receptor D4 (*DRD4*) gene. Moreover, Auerbach et al. (2005, 2008, 2010) demonstrated that a family history was predictive of early temperamental characteristics (analogous to reactivity and regulation considered here) and that it was a combination of family history (in their study designated by father history) and the seven-repeat allele of the *DRD4* gene that predicted ADHD symptoms. Our results suggest that the consideration of specific candidate genes (or gene combinations), perhaps in conjunction with general indicators of a family history of ADHD, may be important for elucidating conditional associations between early temperament and later risk for ADHD.

Similar to family history of ADHD, although there was a main effect of gender on teacher-reported behavior (males > females), there was no evidence that gender moderated the association between early temperament and later ADHD behaviors. In general, there were consistent, albeit modest-sized, gender differences in multiple dimensions of temperament and ADHD behavior. However, we found no evidence that the effects of temperament as a risk factor for ADHD behaviors differed by gender. Child gender was intended to serve as an exploratory moderator. The results of this study do not lend any support to the prospect that the risk for later ADHD behaviors that results from temperament differs for boys or girls. The large sample size, combined with balanced distribution of gender (which is often not evident in clinical samples), lends credence to this conclusion.

Three other features of this study were noteworthy. First, we focused on the prediction of overall levels of ADHD behaviors. We used mean scores instead of symptom counts in order to preserve as much variation in ADHD behavior as possible. In the last 5 years, numerous studies, including our own that involved this sample, have indicated that ADHD symptoms are best characterized by an overall (general) ADHD factor, along with separate inattentive and hyperactive–impulsive factors (Martel, von Eye, & Nigg, 2010; Toplak et al., 2009; Willoughby, Blanton, et al., 2015). Our prediction of an overall ADHD behaviors score represented this general factor. In results that were not presented, we considered the separate prediction of overall ADHD and specific inattentive and hyperactive-impulsive factors. However, the results predicting the general factor were virtually identical to those presented here, while the results predicting the specific factors were largely null and in some cases characterized by peculiarities. Second, regardless of the method that was used (variable or person centered), results were very similar for both parent and teacher reports of ADHD behaviors, even though they were only moderately correlated ($r = .39$; see Table 1). This cross-informant consistency adds credibility to the contributions of early temperament to later ADHD behaviors in multiple contexts. Third, by relying exclusively on observational measures of temperament and parent and teacher ratings of ADHD behaviors, we eliminated shared-method variance as a potential confounder, which was a liability of many previous studies (Lahey, 2004).

Although individual differences in temperament across the first 3 years of life were associated with parent- and teacher-rated ADHD behaviors in first grade, the magnitude of the reported associations was modest. ($R^2 = .10$–.17). Moreover, the time and effort costs that were associated with the collection and coding of these temperament data were substantial. In our opinion, the collection of early objective measures of temperament as a means of identifying risk for future ADHD may be most useful in research settings. For example, this information might be used to enroll young children into prevention or early intervention programs that are intended to reduce the risk for later ADHD. However, it is unlikely that temperamental data would be clinically valuable at the level of individual children. Even with the elevated rates of combined parent- and teacher-reported ADHD symptoms that were evident in this sample, the majority of children in each temperamental profile group (including those presumed to be risks for ADHD) did not meet diagnostic criteria for ADHD. The ability to predict a diagnosis of ADHD at school age from data that were collected early in life remains an elusive goal for ADHD, as well as a host of other psychiatric or medical disorders (e.g., autism). It seems likely that these efforts will require consideration of a broad range of risk and protective factors (e.g., temperament and genetic risk) as well as early developmental processes (e.g., socialization efforts) in order to be fruitful.

This study was characterized by at least four limitations. First, we represented the constructs of activity level, fear and anger reactivity, and regulation as individual difference variables, using multiple indicators that were obtained across multiple home visits (6–24 months for reactivity and 15–36 months for regulation). This approach fails to acknowledge developmental changes in temperament that may be important for the prediction of later ADHD behaviors. Second, this study focused on the additive and multiplicative effects of temperament constructs, including as a function of a family history of ADHD and child gender. However, we did not consider contextual factors (e.g., characteristics of the household and/or caregiving environment) that may moderate temperament risk and subsequent ADHD. A systematic investigation of potential moderators between early temperament and later ADHD is an important direction for future research. Third, although we conceptualized temperament from the perspective of reactivity and regulation, our measurement of activity level is not equivalent to positive exuberance, for which we did not have good measurement and which may be relevant to ADHD (see, e.g., Karalunas et al., 2014). Similarly, our regulatory construct was dominated by measures of "cool" executive function (inhibitory control and attention shifting) and attentional persistence. Consideration of "hot" tasks (Zelazo & Carlson, 2012), including delay of gratification or delay aversion, may be preferable for the prediction of ADHD. Fourth, we predicted individual differences in parent- and teacher-rated ADHD behaviors, not clinical diagnosis.

In sum, despite long-standing interest in the conceptual and empirical overlap between child temperament and subsequent risk for ADHD, relatively few studies have utilized longitudinal data to establish prospective associations between temperament that is measured in the first years of life and risk for ADHD in middle childhood. This study provides some of the strongest evidence to date that early emerging individual differences in temperamental reactivity and regulation are associated with increased risk for elevations in school-age ADHD behaviors. However, the magnitude of these associations was modest. An important direction for future research involves the identification of other attributes of children, as well as their social ecologies, that differentially contribute to an improved understanding of which children are most and least likely to be diagnosed with ADHD in middle childhood. This knowledge would facilitate early intervention and is consistent with the precision medicine initiative that will transform physical and mental health care as we know it.

References

Altink, M. E., Arias-Vasquez, A., Franke, B., Slaats-Willemse, D. I. E., Buschgens, C. J. M., Rommelse, N. N. J., et al. (2008). The dopamine receptor D4 7-repeat allele and prenatal smoking in ADHD-affected children and their unaffected siblings: No gene–environment interaction. *Journal of Child Psychology and Psychiatry*, *49*, 1053–1060. doi:10.1111/j.1469-7610.2008.01998.x

Auerbach, J. G., Atzaba-Poria, N., Berger, A., Landau, R., Arbelle, S., Raz, Y., et al. (2010). Dopamine risk and paternal ADHD symptomatology associated with ADHD symptoms in four and a half-year-old boys. *Psychiatric Genetics*, *20*, 160–165.

Auerbach, J. G., Berger, A., Atzaba-Poria, N., Arbelle, S., Cypin, N., Friedman, A., et al. (2008). Temperament at 7, 12, and 25 months in children at familial risk for ADHD. *Infant and Child Development*, *17*, 321–338.

Auerbach, J. G., Landau, R., Berger, A., Arbelle, S., Faroy, M., & Karplus, M. (2005). Neonatal behavior of infants at familial risk for ADHD. *Infant Behavior and Development*, *28*, 220–224.

Bayley, N. (1969). *Bayley Scales of Infant Development*. New York: Psychological Corp.

Bell, L., Kellison, I., Garvan, C. W., & Bussing, R. (2010). Relationships between child-reported activity level and task orientation and parental attention-deficit/hyperactivity disorder symptom ratings. *Journal of Developmental and Behavioral Pediatrics*, *31*, 233–237.

Berger, A., Alyagon, U., Hadaya, H., Atzaba-Poria, N., & Auerbach, J. G. (2013). Response inhibition in preschoolers at familial risk for attention deficit hyperactivity disorder: A behavioral and electrophysiological stop-signal study. *Child Development*, *84*, 1616–1632. doi:10.1111/Cdev.12072

Biederman, J., & Faraone, S. V. (2004). The Massachusetts General Hospital studies of gender influences on attention-deficit/hyperactivity disorder in youth and relatives. *Psychiatric Clinics of North America*, *27*, 225.

Biederman, J., Faraone, S. V., Keenan, K., Benjamin, J., Krifcher, B., Moore, C., & et al. (1992). Further evidence for family-genetic risk factors in attention deficit hyperactivity disorder: Patterns of comorbidity in probands and relatives in psychiatrically and pediatrically referred samples. *Archives of General Psychiatry*, *49*, 728–738.

Bussing, R., Lehninger, F., & Eyberg, S. (2006). Difficult child temperament and attention-deficit/hyperactivity disorder in preschool children. *Infants and Young Children*, *19*, 123–131.

Carlson, S. M., Mandell, D. J., & Williams, L. (2004). Executive function and theory of mind: Stability and prediction from ages 2 to 3. *Developmental Psychology*, *40*, 1105–1122.

Das Banerjee, T., Middleton, F., & Faraone, S. V. (2007). Environmental risk factors for attention-deficit hyperactivity disorder. *Acta Paediatrica*, *96*, 1269–1274.

Else-Quest, N. M., Hyde, J. S., Goldsmith, H. H., & Van Hulle, C. A. (2006). Gender differences in temperament: A meta-analysis. *Psychological Bulletin*, *132*, 33–72.

Epstein, J. N., Conners, C. K., Erhardt, D., Arnold, L. E., Hechtman, L., Hinshaw, S. P., et al. (2000). Familial aggregation of ADHD characteristics. *Journal of Abnormal Child Psychology*, *28*, 585–594.

Erskine, H. E., Ferrari, A. J., Nelson, P., Polanczyk, G. V., Flaxman, A. D., Vos, T., et al. (2013). Research Review: Epidemiological modelling of attention-deficit/hyperactivity disorder and conduct disorder for the Global Burden of Disease Study 2010. *Journal of Child Psychology and Psychiatry*, *54*, 1263–1274. doi:10.1111/Jcpp.12144

Fischer, M., Barkley, R. A., Smallish, L., & Fletcher, K. (2002). Young adult follow-up of hyperactive children: Self-reported psychiatric disorders, comorbidity, and the role of childhood conduct problems and teen CD. *Journal of Abnormal Child Psychology*, *30*, 463–475.

Fischer, M., Barkley, R. A., Smallish, L., & Fletcher, K. (2007). Hyperactive children as young adults: Driving abilities, safe driving behavior, and adverse driving outcomes. *Accident Analysis and Prevention*, *39*, 94–105. doi:10.1016/j.aap.2006.06.008

Foley, M., McClowry, S. G., & Castellanos, F. X. (2008). The relationship between attention deficit hyperactivity disorder and child temperament. *Journal of Applied Developmental Psychology*, *29*, 157–169.

Froehlich, T. E., Anixt, J. S., Loe, I. M., Chirdkiatgumchai, V., Kuan, L., & Gilman, R. C. (2011). Update on environmental risk factors for attention-deficit/hyperactivity disorder. *Current Psychiatry Reports*, *13*, 333–344. doi:10.1007/s11920-011-0221-3

Froehlich, T. E., Lanphear, B. P., Dietrich, K. N., Cory-Slechta, D. A., Wang, N., & Kahn, R. S. (2007). Interactive effects of a DRD4 polymorphism, lead, and sex on executive functions in children. *Biological Psychiatry*, *62*, 243–249.

Galera, C., Cote, S. M., Bouvard, M. P., Pingault, J. B., Melchior, M., Michel, G., et al. (2011). Early risk factors for hyperactivity-impulsivity and inattention trajectories from age 17 months to 8 years. *Archives of General Psychiatry*, *68*, 1267–1275.

Gerardi-Caulton, G. (2000). Sensitivity to spatial conflict and the development of self-regulation in children 24-36 months of age. *Developmental Science*, *3*, 397–404.

Goldsmith, H., & Rothbart, M. (1996). *The Laboratory Temperament Assessment Battery: Prelocomotor version. 3.0*. Madison, WI: University of Wisconsin, Department of Psychology.

Guerin, D. W., Gottfried, A. W., & Thomas, C. W. (1997). Difficult temperament and behaviour problems: A longitudinal study from 1.5 to 12 years. *International Journal of Behavioral Development*, *21*, 71–90.

Gurevitz, M., Geva, R., Varon, M., & Leitner, Y. (2014). Early markers in infants and toddlers for development of ADHD. *Journal of Attention Disorders*, *18*, 14–22. doi:10.1177/1087054712447858

Healey, D. M., Marks, D. J., & Halperin, J. M. (2011). Examining the interplay among negative emotionality, cognitive functioning, and attention-deficit/hyperactivity disorder symptom severity. *Journal of the International Neuropsychological Society*, *17*, 502–510. doi:10.1017/S1355617711000294

Hemmi, M. H., Wolke, D., & Schneider, S. (2011). Associations between problems with crying, sleeping and/or feeding in infancy and long-term behavioural outcomes in childhood: A meta-analysis. *Archives of Disease in Childhood*, *96*, 622–629. doi:10.1136/adc.2010.191312

Hermens, D. F., Williams, L. M., Lazzaro, I., Whitmont, S., Melkonian, D., & Gordon, E. (2004). Sex differences in adult ADHD: A double dissociation in brain activity and autonomic arousal. *Biological Psychology*, *66*, 221–233.

Karalunas, S. L., Fair, D., Musser, E. D., Aykes, K., Iyer, S. P., & Nigg, J. T. (2014). Subtyping attention-deficit/hyperactivity disorder using temperament dimensions toward biologically based nosologic criteria. *JAMA Psychiatry*, *71*, 1015–1024. doi:10.1001/jamapsychiatry.2014.763

Klein, R. G., Mannuzza, S., Olazagasti, M. A. R., Roizen, E., Hutchison, J. A., Lashua, E. C., et al. (2012). Clinical and functional outcome of childhood attention-deficit/hyperactivity disorder 33 years later. *Archives of General Psychiatry*, *69*, 1295–1303. doi:10.1001/archgenpsychiatry.2012.271

Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in early childhood: Continuity and change, antecedents, and implications for social development. *Developmental Psychology*, *36*, 220–232.

Kupfer, D. J., Baltimore, R. S., Berry, D. A., Breslau, N., Ellinwood, E. H., Ferre, J., et al. (2000). National Institutes of Health Consensus Development Conference Statement: Diagnosis and treatment of attention-deficit/hyperactivity disorder (ADHD). *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*, 182–193.

Lahey, B. (2004). Commentary: Role of temperament in developmental models of psychopathology. *Journal of Clinical Child and Adolescent Psychology*, *33*, 88–93.

Lambert, N. M. (1982). Temperament profiles of hyperactive children. *American Journal of Orthopsychiatry*, *52*, 458–467.

Landau, R., Sadeh, A., Vassoly, P., Berger, A., Atzaba-Poria, N., & Auerbach, J. G. (2010). Sleep patterns of 7-week-old infants at familial risk for attention deficit hyperactivity disorder. *Infant Mental Health Journal*, *31*, 630–646.

Langley, K., Heron, J., Smith, G. D., & Thapar, A. (2012). Maternal and paternal smoking during pregnancy and risk of ADHD symptoms in offspring: Testing for intrauterine effects. *American Journal of Epidemiology*, *176*, 261–268. doi:10.1093/Aje/Kwr510

Levy, F., Hay, D. A., Bennett, K. S., & McStephen, M. (2005). Gender differences in ADHD subtype comorbidity. *Journal of the American Academy of Child & Adolescent Psychiatry*, *44*, 368–376.

Linnet, K. M., Dalsgaard, S., Obel, C., Wisborg, K., Henriksen, T. B., Rodriguez, A., et al. (2003). Maternal lifestyle factors in pregnancy risk of attention deficit hyperactivity disorder and associated behaviors: Review of the current evidence. *American Journal of Psychiatry*, *160*, 1028–1040.

Martel, M. M. (2013). Individual differences in attention deficit hyperactivity disorder symptoms and associated executive dysfunction and traits: Sex, ethnicity, and family income. *American Journal of Orthopsychiatry*, *83*, 165–175. doi:10.1111/Ajop.12034

Martel, M. M., Gremillion, M. L., & Roberts, B. (2012). Temperament and common disruptive behavior problems in preschool. *Personality and Individual Differences*, *53*, 874–879. doi:10.1016/j.paid.2012.07.011

Martel, M. M., Lucia, V. C., Nigg, J. T., & Breslau, N. (2007). Sex differences in the pathway from low birth weight to inattention/hyperactivity. *Journal of Abnormal Child Psychology*, *35*, 87–96.

Martel, M. M., & Nigg, J. T. (2006). Child ADHD and personality/temperament traits of reactive and effortful control, resiliency, and emotionality. *Journal of Child Psychology and Psychiatry*, *47*, 1175–1183. doi:10.1111/j.1469-7610.2006.01629.x

Martel, M. M., Nigg, J. T., & Lucas, R. E. (2008). Trait mechanisms in youth with and without attention-deficit/hyperactivity disorder. *Journal of Research in Personality*, *42*, 895–913.

Martel, M. M., Nigg, J. T., & von Eye, A. (2009). How do trait dimensions map onto ADHD symptom domains? *Journal of Abnormal Child Psychology*, *37*, 337–348.

Martel, M. M., von Eye, A., & Nigg, J. T. (2010). Revisiting the latent structure of ADHD: Is there a “g” factor? *Journal of Child Psychology and Psychiatry*, *51*, 905–914.

McIntosh, D. E., & ColeLove, A. S. (1996). Profile comparisons between ADHD and non-ADHD children on the temperament assessment battery for children. *Journal of Psychoeducational Assessment*, *14*, 362–372.

Muris, P., & Ollendick, T. H. (2005). The role of temperament in the etiology of child psychopathology. *Clinical Child and Family Psychology Review*, *8*, 271–289. doi:10.1007/s10567-005-8809-y

Neuman, R. J., Lobos, E., Reich, W., Henderson, C. A., Sun, L. W., & Todd, R. D. (2007). Prenatal smoking exposure and dopaminergic genotypes interact to cause a severe ADHD subtype. *Biological Psychiatry*, *61*, 1320–1328. doi:10.1016/j.biopsych.2006.08.049

Nigg, J. T. (2006). Temperament and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *47*, 395–422.

Nigg, J. T., Goldsmith, H. H., & Sachek, J. (2004). Temperament and attention deficit hyperactivity disorder: The development of a multiple pathway model. *Journal of Clinical Child and Adolescent Psychology*, *33*, 42–53.

Nigg, J. T., Nikolas, M., & Burt, S. A. (2010). Measured gene-by-environment interaction in relation to attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*, 863–873. doi:10.1016/j.jaac.2010.01.025

Nylund, K. L., Asparouhov, T., & Muthen, B. O. (2008). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, *15*, 182. doi:10.1080/10705510701793320

Oberklaid, F., Sanson, A., Pedlow, R., & Prior, M. (1993). Predicting preschool behavior problems from temperament and other variables in infancy. *Pediatrics*, *91*, 113–120.

Pauli-Pott, U., Dalir, S., Mingebach, T., Roller, A., & Becker, K. (2013). Do different ADHD-related etiological risks involve specific neuropsychological pathways? An analysis of mediation processes by inhibitory control and delay aversion. *Journal of Child Psychology and Psychiatry*, *54*, 800–809. doi:10.1111/jcpp.12059

Pelham, W. E., Evans, S. W., Gnagy, E. M., & Greenslade, K. E. (1992). Teacher ratings of DSM-III-R symptoms for the disruptive behavior disorders: Prevalence, factor analyses, and conditional probabilities in a special education sample. *School Psychology Review*, *21*, 285–299.

Pelham, W. E., Foster, E. M., & Robb, J. A. (2007). The economic impact of attention-deficit/hyperactivity disorder in children and adolescents. *Journal of Pediatric Psychology*, *32*, 711–727.

Polanczyk, G., de Lima, M. S., Horta, B. L., Biederman, J., & Rohde, L. A. (2007). The worldwide prevalence of ADHD: A systematic review and metaregression analysis. *American Journal of Psychiatry*, *164*, 942–948.

Ramtekkar, U. P., Reiersen, A. M., Todorov, A. A., & Todd, R. D. (2010). Sex and age differences in attention-deficit/hyperactivity disorder symptoms and diagnoses: Implications for DSM-V and ICD-11. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*, 217–228.

Rettew, D. C., Copeland, W., Stranger, C., & Hudziack, J. J. (2004). Associations between temperament and DSM-IV externalizing disorders in children and adolescents. *Journal of Developmental & Behavioral Pediatrics*, *25*, 383–392.

Rucklidge, J. J. (2010). Gender differences in attention-deficit/hyperactivity disorder. *Psychiatric Clinics of North America*, *33*, 357.

Sagiv, S. K., Epstein, J. N., Bellinger, D. C., & Korrick, S. A. (2013). Pre- and postnatal risk factors for ADHD in a nonclinical pediatric population. *Journal of Attention Disorders*, *17*, 47–57. doi:10.1177/1087054711427563

Seidman, L. J., Biederman, J., Faraone, S. V., Milberger, S., Norman, D., Seiverd, K., et al. (1995). Effects of family history and comorbidity on the neuropsychological performance of children with ADHD: Preliminary findings. *Journal of the American Academy of Child & Adolescent Psychiatry*, *34*, 1015–1024. doi:10.1097/00004583-199508000-00011

Seidman, L. J., Biederman, J., Monuteaux, M. C., Valera, E., Doyle, A. E., & Faraone, S. V. (2005). Impact of gender and age on executive functioning: Do girls and boys with and without attention deficit hyperactivity disorder differ neuropsychologically in preteen and teenage years? *Developmental Neuropsychology*, *27*, 79–105.

Skoglund, C., Chen, Q., D’Onofrio, B. M., Lichtenstein, P., & Larsson, H. (2014). Familial confounding of the association between maternal smoking during pregnancy and ADHD in offspring. *Journal of Child Psychology and Psychiatry*, *55*, 61–68. doi:10.1111/Jcpp.12124

Sonuga-Barke, E. J. S., & Halperin, J. M. (2010). Developmental phenotypes and causal pathways in attention deficit/hyperactivity disorder: Potential targets for early intervention? *Journal of Child Psychology and Psychiatry*, *51*, 368–389.

Stawicki, J. A., Nigg, J. T., & von Eye, A. (2006). Family psychiatric history evidence on the nosological relations of DSM-IV ADHD combined and inattentive subtypes: New data and meta-analysis. *Journal of Child Psychology and Psychiatry*, *47*, 935–945. doi:10.1111/j.1469-7610.2006.01628.x

Stifter, C. A., & Braungart, J. M. (1995). The regulation of negative reactivity in infancy—Function and development. *Developmental Psychology*, *31*, 448–455. doi:10.1037/0012-1649.31.3.448

Stifter, C. A., & Corey, J. M. (2001). Vagal regulation and observed social behavior in infancy. *Social Development*, *10*, 189–201.

Stifter, C. A., & Dollar, J. (in press). Temperament and developmental psychopathology. In D. Cicchetti (Ed.), *Developmental psychopathology* (Vol. 3, 3rd ed.). New York: Wiley.

Sullivan, E. L., Holton, K. F., Nousen, E. K., Barling, A. N., Sullivan, C. A., Propper, C. B., et al. (2015). Early identification of ADHD risk via infant temperament and emotion regulation: A pilot study. *Journal of Child Psychology and Psychiatry*, *56*, 949–957. doi:10.1111/jcpp.12426

Thapar, A., Cooper, M., Eyre, O., & Langley, K. (2013). Practitioner Review: What have we learnt about the causes of ADHD? *Journal of Child Psychology and Psychiatry*, *54*, 3–16. doi:10.1111/j.1469-7610.2012.02611.x

Thapar, A., Fowler, T., Rice, F., Scourfield, J., van den Bree, M., Thomas, H., et al. (2003). Maternal smoking during pregnancy and attention deficit hyperactivity disorder symptoms in offspring. *American Journal of Psychiatry*, *160*, 1985–1989. doi:10.1176/appi.ajp.160.11.1985

Thapar, A., Rice, F., Hay, D., Boivin, J., Langley, K., van den Bree, M., et al. (2009). Prenatal smoking might not cause attention-deficit/hyperactivity disorder: Evidence from a novel design. *Biological Psychiatry*, *66*, 722–727. doi:10.1016/j.biopsych.2009.05.032

Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Bruner/Mazel.

Tofighi, D., & Enders, C. K. (2007). Identifying the correct number of classes in growth mixture models. In G. R. Hancock & K. M. Samuelsen (Eds.), *Advances in latent variable mixture models* (pp. 317–341). Greenwich, CT: Information Age.

Toplak, M. E., Pitch, A., Flora, D. B., Iwenofu, L., Ghelani, K., Jain, U., et al. (2009). The unity and diversity of inattention and hyperactivity/impulsivity in ADHD: Evidence for a general factor with separable dimensions. *Journal of Abnormal Child Psychology*, *37*, 1137–1150. doi:10.1007/s10802-009-9336-y

Vernon-Feagans, L., Cox, M., & the Family Life Project Investigators. (2013). The Family Life Project: An epidemiological and developmental study of young children living in poor rural communities. *Monographs of the Society for Research in Child Development*, *78*, 1–150.

Willcutt, E. G., Nigg, J. T., Pennington, B. F., Solanto, M. V., Rohde, L. A., Tannock, R., et al. (2012). Validity of DSM-IV attention deficit/hyperactivity disorder symptom dimensions and subtypes. *Journal of Abnormal Psychology*.

Willoughby, M. T., Blair, C. B., Wirth, R. J., Greenberg, M., & the Family Life Project Investigators. (2010). The measurement of executive function at age 3 years: Psychometric properties and criterion validity of a new battery of tasks. *Psychological Assessment*, *22*, 306–317.

Willoughby, M. T., Blanton, Z. E., & the Family Life Project Investigators. (2015). Replication and external validation of a bi-factor parameterization of attention deficit/hyperactivity symptomatology. *Journal of Clinical Child and Adolescent Psychology*, *44*, 68–79. doi:10.1080/15374416.2013.850702

Willoughby, M. T., Pek, J., Greenberg, M. T., & the Family Life Project Investigators. (2012). Parent-reported attention deficit/hyperactivity symptomatology in preschool-aged children: Factor structure, developmental change, and early risk factors. *Journal of Abnormal Child Psychology*, *40*, 1301–1312. doi:10.1007/s10802-012-9641-8

Willoughby, M. T., Stifter, C. A., Gottfredson, N. C., & the Family Life Project Investigators. (2014). The epidemiology of observed temperament: Factor structure and demographic group differences. *Infant Behavior & Development*. Advance online publication.

Willoughby, M. T., Stifter, C. A., Gottfredson, N. C., & the Family Life Project Investigators. (2015). The epidemiology of observed temperament: Factor structure and demographic group differences. *Infant Behavior & Development*, *39*, 21–34. doi:10.1016/j.infbeh.2015.02.001

Yoshimasu, K., Barbaresi, W. J., Colligan, R. C., Killian, J. M., Voigt, R. G., Weaver, A. L., et al. (2010). Gender, attention-deficit/hyperactivity disorder, and reading disability in a population-based birth cohort. *Pediatrics*, *126*, E788–E795. doi:10.1542/peds.2010-1187

Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, *6*, 354–360. doi:10.1111/j.1750-8606.2012.00246.x

Appendix A

Description of indicators of temperament factor scores

Reactivity: home visitor ratings (6–24 months). After each home visit, both home visitors independently made global ratings of children's behavior (see Stifter & Corey, 2001, for precedent) using items that were adapted from the Infant Behavior Record (Bayley, 1969). Hence, at the 6- and 24-month assessments, four independent ratings (two ratings per visit for two visits) were available. At the 15-month assessment, two independent ratings (two rating for the single home visit) were available. Each item was rated on a 9-point Likert scale.

Home visitor ratings of the *amount of gross bodily movement* item (Likert anchors: 1 = *stays quietly in one place, with practically no self-initiated movement* to 9 = *hyperactive, cannot be quieted for sedentary tasks*) were used as an indicator of activity level at the 6-, 15-, and 24-month assessments (αs = 0.79, 0.74, and 0.80, respectively). Home visitor ratings of the *reaction to the new or strange* item (Likert anchors: 1 = *accepts the entire situation with no evidence of fear, caution or inhibition of actions* to 9 = *strong indication of fear of the strange, to the extent that he cannot be brought to play or respond to the examiner or tasks*) were used as an indicator of fear reactivity at the 15- and 24-month assessments (αs = 0.74 and 0.75, respectively). Home visitor ratings of the *irritability* item (Likert anchors: 1 = *no irritability, infant passively responses to all stimulation* to 9 = *irritable to all degrees of stimulation encountered throughout the home visit*) were used as an indicator of anger reactivity at the 15- and 24-month assessments (αs = 0.79 and 0.78, respectively).

Reactivity: Observed activity level (6 months). Parents and children participated in a free-play interaction at the 6-month assessment, and videotaped interactions were coded to assess multiple dimensions of global parenting and child behaviors across each interaction. The current study made use of the child activity level code from the 6-month visits. Ratings for activity level (extent to which the child is motorically active during the observation) were made on a 5-point scale ranging from 1 = *not at all characteristic* to 5 = *highly characteristic*. Coders were trained and certified as reliable prior to coding. A minimum of 30% of all observations were double coded throughout the coding period, and discrepancies in coding were resolved by conferencing. Coding pairs exhibited acceptable interrater reliability for the activity code at 6 months (intraclass correlation = .63).

Reactivity: Challenge tasks (15 and 24 months). Infants participated in two tasks that were drawn from the Lab-TAB (Goldsmith & Rothbart, 1996). Each task was videotaped and coded offline by trained research assistants. The *mask task*, which was designed to elicit individual differences in fear reactivity, was administered at the 15- and 24-month home visits. During the task, the home visitor would put on an unusual mask and move her head slowly from side to side while calling the child's name. A total of four masks were presented, for approximately 10 s each. The *toy removal task*, which was designed to elicit anger reactivity, was administered at the 15- and 24-month assessments (Stifter & Braungart, 1995). During the task, the infant and mother play with an interesting toy for 90 s after which the mother removes the toy and places it on her seat. The mother then moves away to converse with the home visitor for 2 min. The mother than returns the toy and resumes speaking with the home visitor. After 1 min, the mother returns and resumes interacting with the child or soothes him/her if needed.

Both tasks were subjected to second-by-second coding using the Better Coding Approach software (Danville, Pennsylvania). Three levels of task-related negative reactivity were coded: low negative reactivity, moderate negative reactivity, and high negative reactivity. A weighted negative reactivity composite was calculated for each task (i.e., the average proportion of task time spent exhibiting *low* = 1, *moderate* = 2, and *high* = 3 levels of reactivity). All coders were trained to achieve at least 0.75 (Cohen κ) reliability on the reactivity coding. Subsequent interrater reliability was calculated on 15% of cases using κcoefficients (κs = 0.96 and 0.93 for the mask task at 15 and 24 months; κs = 0.88 and 0.90 for the toy removal task at 15 and 24 months).

Regulation: Home visitor ratings (24 and 36 months). Home visitor ratings on the *tendency to persist in attention to any one objective, person, or activity, aside from attaining a goal* (Likert anchors: 1 = *fleeting attention span to* 9 = *long-continued absorption in a toy, activity, or person*) item were used as an indicator of regulation at the 24-month assessment (α = 0.75) and on the *task persistence—degree of on-task behavior, persistence in the face of frustration* item (Likert anchors: 1 = *does not stay on task, short attention span to* 9 = *shows high level of task persistence*) was used as an indicator at the 36-month assessment (α = 0.80).

Regulation: Observed persistence (24 and 36 months). Parents and children participated in a joint puzzle-solving interaction at the 24- and 36-month assessments. Videotaped interactions were coded to assess multiple dimensions of global parenting and child behaviors across each interaction. The child persistence codes from the 24- and 36-month visits were examined in the present study. Ratings for child persistence were made on a 7-point scale ranging from 1 = *very low* (e.g., child actively tries to avoid the task and spends as little time as s/he can get away with doing the tasks at all) to 7 = *very high* (e.g., child is persistent and works at each task with an apparent goal of getting correct solutions until the problem is solved or exhaustively approximated). The persistence code was intended to reflect the child problem-solving efforts regardless and independent of the degree to which the parent was instrumental in facilitating the child's persistence. Coders double coded 30% of all observations and exhibited acceptable interrater reliability for the persistence code at the 24- and 36-month assessments (intraclass correlations = .75 and .76, respectively).

Regulation: Direct assessments of inhibitory control and attention shifting (24 and 36 months). Children completed one inhibitory control and one attention shifting task at the 24-month visit. The *snack delay task*, adopted from Kochanska, Murray, and Harlan (2000), was used as an indicator of inhibitory control (delay of gratification). In this task, the experimenter placed a desirable snack (small candy or cracker) underneath a transparent container. The child was told to wait until the experimenter rang a bell before retrieving the snack. Each child completed four trials: a 10-s delay followed by 20-, 30-, and 15-s delay trials. Each response was coded as *no wait* (0), *partial wait* (1), or *full wait* (2), and the mean score across trials indicated delay of gratification. The *reverse categorization task*, adopted from Carlson, Mandell, and Williams (2004), was used as an indicator of attention shifting. In this task, children initially sort small- and big-sized blocks into corresponding-sized buckets but are subsequently asked to sort small blocks into big buckets and vice versa. Three trials (practice, preswitch, and switch) of six blocks were administered. Children were designated as passing a trial if they correctly sorted four of six blocks. The sum of passed trials (0–3) was the dependent variable.

Children completed three inhibitory control and one attention shifting task at the 36-month assessment. Full task descriptions, administration and scoring procedures, and psychometric properties of individual tasks were elaborated elsewhere (Willoughby, Blair, Wirth, Greenberg, & the Family Life Project Investigators, 2010). Briefly, the silly sounds Stroop task (inhibitory control) asks children to make a barking sound when shown pictures of a cat and a meow sound when shown pictures of a dog. The spatial conflict task (inhibitory control) is a Simon task modeled after Gerardi-Caulton (2000). The animal go/no-go task (inhibitory control) presents a series of pictures of animals to children and asks that they click a button every time that they see an animal (go trials) except when that animal is a pig (no-go trials). The something's the same task (attention shifting) presents children with a page containing two pictures that are similar along one dimension (content, color, or size) and asks them to indicate which of these pictures is similar to a third picture.