



Using Repeated-Measures Data to Make Stronger Tests of the Association between Executive Function Skills and Attention Deficit/Hyperactivity Disorder Symptomatology in Early Childhood

Michael T. Willoughby¹ · Amanda C. Wylie¹ · Clancy B. Blair²

Published online: 14 May 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Theoretical models of Attention deficit/hyperactivity disorder (ADHD) have long implicated executive function (EF) skills as contributing to the etiology, maintenance, and changes in ADHD symptomatology over time. Although there is interest making within-person inferences (i.e., deficits in EF skills give rise to ADHD behaviors), most of the evidence has been derived from studies that conflated between- and within-person sources of variance. Here, we use repeated-measures data to test within-person association between EF skills and ADHD behaviors. Participants included 1160 children from the Family Life Project, an ongoing prospective longitudinal study of child development in low-income, nonmetropolitan communities. We tested the magnitude of the association between EF skills and ADHD behaviors when children were 3, 4, and 5 years old. Consistent with meta-analyses, unadjusted bivariate associations between EF and ADHD (which reflect combined between- and within-person variation) were of moderate magnitude ($r_s = -0.20$ to -0.30). However, after controlling for all time-invariant, between-person sources of variation, the within-person associations between EF skills and ADHD behaviors were weak ($\beta_s = 0.04$ to -0.05 , $p_s = 0.01$). These results suggest that EF skills may contribute less prominently to ADHD behaviors in early childhood than is commonly assumed and provoke broader questions about developmental models of ADHD.

Keywords Attention deficit/hyperactivity disorder · Executive function · Early childhood · Structural equation modeling · Fixed-effects analysis

Attention deficit/hyperactivity disorder (ADHD) is a childhood-onset disorder that is characterized by developmentally inappropriate levels of hyperactive, impulsive, and inattentive behaviors that cause impairment in multiple settings (American Psychiatric Association 2013). Over the past 4 decades, the most prominent changes to the diagnostic criteria for ADHD have involved the relative emphasis of specific behaviors (inattention, hyperactivity, impulsivity) or the conceptualization of how these behaviors relate to each other. Most recently, researchers have emphasized that a single general factor explains most of the

variation in ADHD symptoms and that this general factor is more reliable than the subdimensions of inattention and hyperactivity-impulsivity (Arias et al. 2018; Gomez et al. 2018). As such, in this study we focus on individual differences in overall levels of ADHD behaviors.

A longstanding belief has been that individual differences in ADHD behaviors reflect neurobiological dysfunction (Levy 1991; Levy et al. 1998; Swanson et al. 1998; Wender 1975; Zimetkin and Rapoport 1987). Although the sophistication of methods has increased over time (e.g., from gross brain abnormalities to differences in neural network organization; from heritability to genome-wide association or epigenetic expression method studies), the underlying assumptions about the origins of ADHD behaviors remain unchanged (Castellanos and Aoki 2016; Cortese 2012; Middeldorp et al. 2016; Walton et al. 2017). ADHD behaviors are presumed to reflect cognitive impairments, which reflect aberrant neurobiological development, which results from complex interactions between genes and environmental factors.

Decades of research have investigated the specific cognitive processes that are presumed to give rise to ADHD

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10802-019-00559-w>) contains supplementary material, which is available to authorized users.

✉ Michael T. Willoughby
mwilloughby@rti.org

¹ RTI International, 3040 Cornwallis Road, Research Triangle Park, NC 27709, USA

² New York University, New York, NY, USA

behaviors. Although early conceptualizations of ADHD as a purely attention disorder have been dismissed (Huang-Pollock et al. 2005), numerous cognitive processes—including executive functions (EFs), delay aversion, reaction-time variability, and vigilance/arousal—have been linked to ADHD (Coghill et al. 2014; Sonuga-Barke et al. 2010; Tamm et al. 2012). The current consensus is that ADHD is a neuropsychologically heterogeneous disorder across the lifespan (e.g., Kofler et al. 2019; Mostert et al. 2015; Sjowall and Thorell 2019; Wahlstedt et al. 2009). That is, as a group, individuals with ADHD exhibit impairments in multiple cognitive domains relative to their typically developing peers. However, the specific cognitive impairments vary across individuals with ADHD, no single cognitive domain is necessary or sufficient to inform diagnosis, and a minority of individuals with ADHD do not appear to exhibit any cognitive impairment.

Meta-analytic studies have consistently reported medium-sized associations between multiple cognitive processes and ADHD symptomatology in middle childhood through adulthood (Alderson et al. 2007; Doyle 2006; Hervey et al. 2004; Homack and Riccio 2004; Willcutt et al. 2005). This evidence has only been extended to early childhood relatively recently (Pauli-Pott and Becker 2011; Schoemaker et al. 2013). In their meta-analysis of 23 studies, which included over 3000 children, Pauli-Pott and Becker (2011) reported moderate-sized associations ($r_s = .18-.38$) between multiple domains of cognitive function (i.e., response suppression, interference, delay aversion, vigilance, working memory) and ADHD symptomatology. In their meta-analysis of 22 studies, which included over 4000 children, Schoemaker et al. (2013) reported a moderate-sized association ($r = .22$) between EF skills and multiple domains of externalizing behavior, including ADHD ($r_s = 0.24, 0.17, 0.13$ for inhibitory control, working memory, and cognitive flexibility, respectively). Based on this evidence, researchers have proposed early intervention efforts that target improved cognitive function as a strategy to prevent ADHD or mitigate ADHD-related impairments (Sonuga-Barke and Halperin 2010).

The rationale for early interventions that target improvements in children's cognitive function as a strategy for reducing ADHD behaviors is predicated on the longstanding notion that ADHD behaviors result from cognitive dysfunction (i.e., these are *within-person* associations). However, this assumption is not well founded. Meta-analyses report group differences between ADHD and non-ADHD youth on (or relate continuous measures of ADHD behaviors to) cognitive performance measures. However, the individual studies from which effect sizes are derived primarily involve cross-sectional or case-control designs, which conflate between- and within-person sources of variation. Even when the studies that contribute to meta-analyses involve longitudinal designs, the descriptive statistics that are used to generate effect sizes typically conflate between- and within-person sources of

variation. Hence, the interpretation of an average effect size of EF task performance and ADHD behaviors of $r = 0.30$ is ambiguous. Children with higher levels of ADHD behaviors may perform more poorly on EF tasks because EF deficits contribute to their ADHD behaviors. Alternatively, children with higher levels of ADHD behaviors may perform more poorly on EF tasks because of other differences that co-occur with ADHD behaviors (e.g., unmeasured genetic, cognitive, or experiential factors). The former interpretation reflects a within-person association, while the latter reflects a between-person association. With few exceptions (e.g., Arnett et al. 2016), studies of ADHD have rarely explicitly attended to between versus within-person associations.

Measuring both EF skills and ADHD behaviors in the same individuals across multiple occasions is one way in which between- and within-person sources of variation can be disaggregated. However, this disaggregation is contingent on how the repeated-measures data are analyzed (Curran and Bauer 2011). Conceptually, one can imagine a set of time-invariant variables that contribute to increased levels of ADHD behaviors or that equivalently increased risk for diagnosis (e.g., parental ADHD, low birth weight). These time-invariant risk factors are assumed to exert a constant effect on ADHD behaviors (i.e., on any given measurement occasion, children who were exposed to time-invariant risk factors will, on average, exhibit higher levels of ADHD behaviors) and represent between-person sources of variation. Importantly, we do not have to specify what these time-invariant variables are (or even measure them) in order to estimate their effects. If repeated measures of ADHD are available, we can estimate the contributions of all measured and unmeasured time-invariant contributions to ADHD behaviors based on the assumption that they exert a constant effect at each measurement occasion. At the same time, one can imagine a set of time-varying variables that contribute to increased levels of ADHD behaviors (e.g., EF skills). In contrast to time-invariant risk factors, time-varying risk factors change over time and can exert time-dependent contributions to ADHD behaviors. Time-varying risk factors must be explicitly measured (their impacts cannot be inferred) and they represent within-person associations (i.e., they reflect the unique association of a risk factor on ADHD at each measurement occasion, above and beyond the influence of all time-invariant risk factors). In the broader social science literature, random-effect and fixed-effect regression models are the two most commonly employed strategies for distinguishing between- and within-person associations using repeated-measures data (Allison 2009; Halaby 2004; Wooldridge 2002). Here, we followed Bollen and Brand's (2010) structural equation modeling approach for testing between and within-person associations between EF skills and ADHD behaviors. Their approach subsumes and expands on traditional random and fixed-effects regression models.

The overarching objective of this study is to explicitly test the between and within-person associations between children's performance on EF skills and ADHD behaviors in early childhood. EF skills are a set of domain general cognitive processes that include inhibitory control, working memory, and cognitive flexibility (Diamond 2013; Garon et al. 2008). EF skills represent one of multiple cognitive processes implicated in ADHD. In early childhood, EF skills are unidimensional (Espy 2017; Karr et al. 2018). Given our focus on early childhood, we prioritized tests of overall EF skills on ADHD behaviors. However, given that inhibitory control has long been considered a defining characteristic of ADHD (Barkley 1997; Nigg 2001), we repeated all analyses focusing on the time-varying effects of inhibitory control on ADHD behaviors. In both cases, our primary interest was determining whether the frequently reported bivariate association between EF skills and ADHD behaviors represented between- versus within-person sources of variation. Based on the meta-analytic results of Pauli-Pott and Becker (2011) and Schoemaker et al. (2013), we hypothesized that the unadjusted bivariate association between EF skills and ADHD behaviors would be $r \approx 0.30$, with a somewhat smaller association at age 3 versus 5 years. We also hypothesized that the within-person association between EF skills and ADHD behaviors would be smaller in magnitude after we controlled for all time-invariant, between-person sources of variation, though the specific size of the within-person effect was unclear given the absence of directly relevant previous studies.

Method

Participants and Procedures

The Family Life Project is a longitudinal study of rural poverty that involves families who delivered a new child between September 2003 and August 2004 in one of six counties in eastern North Carolina or central Pennsylvania. Complex sampling procedures were employed to recruit a representative sample of 1292 children from rural counties, with oversampling of low-income families in both states and of African American families in North Carolina. A full characterization of the sampling plan and study has been detailed elsewhere (Vernon-Feagans et al. 2013). The University of North Carolina at Chapel Hill Internal Review Board provided oversight of the Family Life Project.

Following newborn hospital screening, families who were selected and agreed to participate were formally enrolled into the study by receiving a home visitor when the target child was approximately 2 months old. Participating children and their families subsequently completed multiple home and school visits. The current study makes use of data that were collected in home visits when children were 3 years, 4 years,

and 5 years old. Specifically, this study is limited to 1160 children who had ADHD or EF data from at least one of the age 3-, 4-, or 5-year visits. Participating children ($N = 1160$) did not differ from nonparticipating children ($N = 132$) with respect to state of residence (40% vs. 36% residing in Pennsylvania, $p = 0.51$), living in a household that was recruited into the low-income stratum (78% vs. 73% poor, $p = 0.33$), primary caregiver educational status at study enrollment (80% vs. 82% with a high school degree or general education diploma, $p = 0.62$), sex of the child (51% vs. 54% male, $p = 0.57$), race of the child (43% vs. 36% African American, $p = 0.28$), or child first-born status (39% vs. 46%, $p = 0.25$). The average age of children at the 3-year visit was $M = 3.09$ years ($SD = 0.15$), at the 4-year visit was $M = 4.03$ years ($SD = 0.13$), and at the 5-year visit was $M = 5.05$ years ($SD = 0.27$). The child's primary caregiver completed ADHD ratings at each visit.

Measures

ADHD Symptom Ratings Consistent with previous studies (Pelham et al. 1992), primary caregivers rated the presence of 18 DSM-IV symptoms for ADHD using a 4-point Likert scale (0 = not at all, 1 = just a little, 2 = pretty much, 3 = very much) at the age 3-, 4-, and 5-year visits. We previously demonstrated that the ADHD items in this sample at these ages were most parsimoniously summarized by a single factor (Willoughby et al. 2012b). As such, the mean rating of all 18 items was used to summarize ADHD at each assessment ($\alpha_s = 0.91, 0.91, \text{ and } 0.93$ at ages 3, 4, and 5 years, respectively). Of the 1160 sampled children, 5 (0.4%) did not have an ADHD rating at any timepoint, 72 (6.2%) had a rating at 1 timepoint, 97 (8.4%) had a rating at 2 timepoints, and 986 (85.0%) had a rating at all 3 timepoints.

Executive Function (Willoughby and Blair 2016) A common battery of EF tasks was administered at the age 3-, 4-, and 5-year visits. One research assistant presented EF tasks to the child using an open, spiral-bound flipbook (20 cm \times 36 cm [8 in. \times 14 in.]), and a second recorded the child's responses into a laptop for subsequent scoring. A standard script and administration procedures were employed throughout (i.e., items were demonstrated for each task, and children were given up to three practice trials to demonstrate their ability to understand task demands; children who did not demonstrate understanding during training were not administered the full task). The task administration procedures; the psychometric properties of individual tasks; and the overall battery score, retest reliability, and criterion validity of these tasks have been elaborated elsewhere (Willoughby et al. 2010; Willoughby and Blair 2011; Willoughby et al. 2012a, b, c). Abbreviated task descriptions are provided below.

Working Memory Span (WMS; Working Memory) This span-like task required children to perform the operation of naming and holding in mind two pieces of information simultaneously (i.e., the names of colors and animals in pictures of “houses”) and to activate one of them (i.e., animal name) while overcoming interference occurring from the other (i.e., color name). Items were more difficult as the number of houses increased (each house included a picture of a color and animal). The WMS was administered at the age 3-, 4-, and 5-year visits and completed by 787 (67.8%), 958 (82.6%), and 983 (84.7%) children at the 3-, 4-, and 5-year visits, respectively.

Pick the Picture Game (PTP; Working Memory) This self-ordered pointing task presented children with a series of 2, 3, 4, or 6 pictures in a set. Children were instructed to pick a picture and on each subsequent page instructed to pick a different picture within the set until each picture had “received a turn.” This task required working memory because children had to update which pictures in each item set they had already touched (the spatial location of pictures changed across trials and was uninformative). Pilot testing had revealed that the PTP was too difficult for many 3-year-olds and therefore was administered only at the 4- and 5-year assessments and completed by 934 (80.5%) and 1004 (86.6%) children at the 4- and 5-year visits, respectively.

Silly Sounds Stroop (SSS; Inhibitory Control) This Stroop task presented children with pictures of cats and dogs and asked children to make the sound opposite of what was associated with each picture (e.g., to meow when showed picture of a dog). This task required inhibitory control, as children had to inhibit the tendency to associate bark and meow sounds with dogs and cats, respectively. The SSS was administered at the age 3-, 4-, and 5-year visits and completed by 479 (41.3%), 894 (77.1%), and 995 (85.8%) children at the 3-, 4-, and 5-year visits, respectively.

Spatial Conflict (SC; Inhibitory Control) This task presented children with a response card that had one picture of a car and another of a boat. Initially, all stimuli (pictures of cars or boats identical to that on the response card) appeared directly above the corresponding photo on the response card (i.e., in a location that was spatially compatible with their placement on the response card). Children were instructed to touch “their” boat or car, according to which one they saw on the stimulus page. Subsequently, test items required a contra-lateral response (i.e., children were to touch their picture of the car even if its position relative to the stimulus card was reversed). This task required inhibitory control as children had to override the spatial location of test stimuli with reference to their response card. The SC was administered at the 3-year assessment and completed by 881 (75.9%) children.

Spatial Conflict Arrows (Inhibitory Control) This task was identical in format to the SC task with the exception that the response card consisted of two black dots (“buttons”) and the test stimuli were arrows, presented one at a time, that pointed to the left or right. Children were instructed to touch the left button for a left-pointing arrow and the right button for a right-pointing arrow. Initially, all left-pointing arrows appeared either centered or directly above the left button, and all right-pointing arrows appeared either centered or above the right button. In the final portion of the assessment, some arrows pointed left but appeared above the right button, or vice versa. The task was administered at the 4- and 5-year assessments and completed by 985 (84.9%) and 1033 (89.1%) children at the 4- and 5-year visits, respectively.

Animal Go no-Go (GNG; Inhibitory Control) This was a standard go no-go task in which children were instructed to click a button (which made an audible sound) every time they saw an animal (i.e., go trials) except when it was a pig (i.e., no-go trials). Varying numbers of go trials appeared before each no-go trial, including, in standard order, 1-go, 3-go, 3-go, 5-go, 1-go, 1-go, and 3-go trials. The GNG was administered at the age 3-, 4-, and 5-year visits and completed by 444 (38.3%), 796 (68.6%), and 980 (84.5%) children at the 3-, 4-, and 5-year visits, respectively.

Something’s the Same Game (STS; Attention Shifting) This task presented children with a pair of pictures for which a single dimension of similarity was noted (i.e., color, size, or object type). Subsequently, a third picture was presented, and children were asked to identify which of the first two pictures was the same as the new picture. This task required the child to shift their attention from the initial labeling to a new dimension of similarity (e.g., a small blue chair and large blue flower would be described as “the same” in color; a small red dog would be described as “the same” as the small blue chair based on size). The STS was administered at the age 3-, 4-, and 5-year visits and completed by 842 (72.6%), 971 (83.7%), and 1024 (88.3%) children at the 3-, 4-, and 5-year visits, respectively.

Executive Function Task Scoring and Composite Formation Children completed an average of 3.0 ($SD = 1.7$) of the 5 tasks that were administered at the age 3 assessment, 4.8 ($SD = 2.1$) of the 6 tasks that were administered at the age 4 assessment, and 5.2 ($SD = 1.9$) of the 6 tasks that were administered at the age 5 assessment. Correlations between the EF tasks within time points were weak to moderate (age 3: $r_s = 0.08$ – 0.25 ; age 4: $r_s = 0.10$ – 0.32 ; age 5: $r_s = 0.07$ – 0.28). As elaborated elsewhere (Willoughby et al. 2012a, b, c), item-response theory models (i.e., graded-response and two-parameter logistic models, depending on the task) were used for task scoring. All tasks exhibited longitudinal measurement invariance and

were scored on a common developmental scale (where $M = 0$, $SD = 1$ was defined as average performance at the age 4 assessment). In our previously published studies, we used individual EF task scores as reflective indicators of a latent variable of EF (e.g., Willoughby et al. 2010; Willoughby et al. 2012a, b, c). However, more recently, we have determined that individual EF task scores are better represented as a composite score (Willoughby et al. 2016, 2017). This EF composite was formed by taking the mean of each child’s task scores and was used as the primary independent variable in the models tested below. Of the 1160 sampled children, 39 (3.4%) did not have an EF composite score at any timepoint, 87 (7.5%) had a composite score at one timepoint, 169 (14.6%) had a composite score at two timepoints, and 865 (74.6%) had a composite score at all three timepoints.

Analytic Strategy

We estimated a series of structural equation models following the approach described by Bollen and Brand (2010). As depicted in Fig. 1, the general form of these models was:

$$ADHD_{it} = \beta_t EF_{it} + \lambda_t \eta_i + \varepsilon_{it}$$

In words, ADHD is a dependent variable for the i^{th} child at time t (here the age 3-, 4-, and 5-year assessments). EF is a time-varying independent variable for the i^{th} child at time t . Beta (β) is a vector of coefficients that relate time-specific measures of EF to ADHD behaviors. Eta sub i (η_i) is a latent variable that represents the impact of all time-invariant factors that influence ADHD behaviors for child i . Lambdas (λ) are coefficients that represent the magnitude of the associations between these latent time-invariant influences on ADHD at time t . One lambda coefficient is fixed to 1 for model identification; we chose age 4 because of our interest in understanding whether the effect of EF

on ADHD differed across age 3 and 5 assessments. The epsilon (ε) term represents the variation in ADHD (for the i^{th} child at time t) that is unexplained by time-varying EF and latent time-invariant factors.

The magnitude and statistical significance of the beta coefficients were of primary substantive interest. However, we began by fitting a sequence of models that imposed increasingly restrictive constraints on the lambda, beta, and epsilon parameters, along with the correlations between eta with EF (represented as phi [φ] parameters). Our sequence of models followed the backward search strategy that was described by Bollen and Brand (2010). We used likelihood ratio tests, global model fit statistics (i.e., comparative fit index [CFI], root mean squared error of approximation [RMSEA]), and information criteria (Akaike information criteria, Bayesian information criteria) to determine which model specification provided the best fit to the observed data. It was only after we had identified this best-fitting model that we interpreted the association between EF and ADHD. All models were fit using a robust full-information maximum-likelihood estimator that took individual probability weights and stratification variables into account to appropriately represent the complex sampling design. Full-information maximum-likelihood estimation also accommodated missing data (i.e., all available EF and ADHD data for each child were included) and is considered a statistical best practice (Schafer and Graham 2002). All models were estimated using version 8.0 of *Mplus* (Muthén & Muthén, 1998–2017).

Results

Bivariate Associations

Descriptive statistics for parent-reported ADHD symptoms and children’s EF composite scores are provided in Table 1.

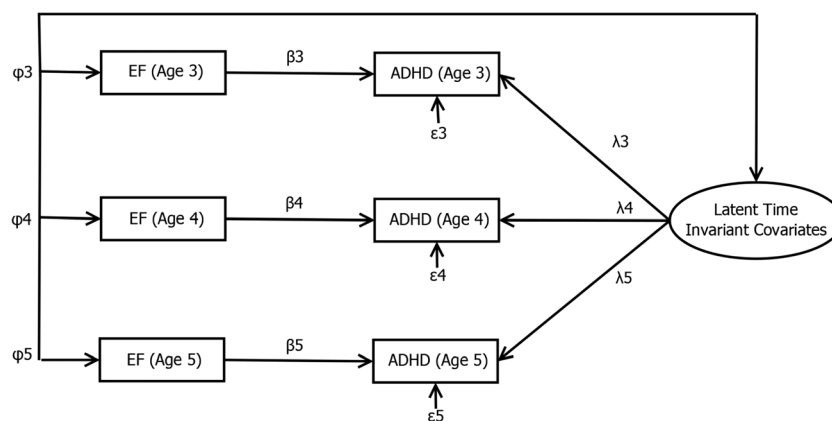


Fig. 1 General Panel Model. *Caption.* ADHD = attention deficit hyperactivity disorder; EF = executive function. β = coefficients relating time-specific measures of EF to ADHD behaviors; ε = variation in ADHD unexplained by time-varying EF and latent time-invariant factors;

λ = coefficients for the magnitude of the associations between latent time-invariant influences on ADHD; φ = correlations between unmeasured time-invariant factors and EF

Table 1 Descriptive statistics

	1	2	3	4	5	6
1. ADHD (age 3)	–					
2. ADHD (age 4)	0.663	–				
3. ADHD (age 5)	0.602	0.724	–			
4. EF (age 3)	–0.205	–0.226	–0.155	–		
5. EF (age 4)	–0.219	–0.302	–0.257	0.385	–	
6. EF (age 5)	–0.240	–0.309	–0.290	0.358	0.606	–
<i>N</i>	1096	1062	1066	973	1009	1038
<i>M</i>	0.998	0.922	0.823	–0.454	–0.033	0.363
<i>SD</i>	0.584	0.579	0.583	0.541	0.507	0.458

ADHD, Attention deficit/hyperactivity disorder; *EF*, executive function; ADHD scores represent mean ratings of 18 symptoms on a 0–3 scale. EF scores represent IRT-based scores that are on a common developmental scale which has a mean of 0 and a standard deviation of 1 based on the age 4 assessment

The across-time correlations for ADHD measures were large ($r_s = 0.60$ – 0.72 , $p_s < 0.001$) and for EF were moderate to large ($r_s = 0.36$ – 0.61 , $p_s < 0.001$). The within-time associations between EF and ADHD were consistent with previous meta-analytic results ($r_s = 0.20$, 0.30 , and 0.29 for the age 3-, 4-, and 5-year assessments, respectively, all $p_s < 0.001$). In terms of mean differences, there were substantial increases in performance on the EF composite and modest decreases in parent-reported ADHD symptomatology.

Model Description, Fit, and Comparisons

A synopsis of global model fit, information criteria, and likelihood ratio tests that were used for model comparisons appears in Table 2. Model 1 was a baseline model in which all parameters were freely estimated and fit the data well, $\chi^2(3) = 3.9$, CFI = 0.99, RMSEA (95% confidence interval [CI]) =

0.016 (0.000, 0.054). Model 2 tested whether the lambda parameters could all be constrained to 1, which would imply that the latent time-invariant influences on ADHD were constant across ages 3–5 years. Although global fit for Model 2 was acceptable (see Table 2), this model fit the data significantly worse than Model 1, $\Delta\chi^2(2) = 8.6$, $p = 0.014$. An inspection of the parameter estimates from Model 1 suggested that the lambda parameter at age 3 was weaker than at ages 4 and 5. Model 2b freed the constraint on the lambda parameter at age 3 and no longer differed in fit from Model 1, $\Delta\chi^2(1) = 2.8$, $p = 0.10$. Building on Model 2b, Model 3 tested whether the beta parameters could be constrained to be equal, which would imply that the time-varying effect of EF on ADHD was constant across ages 3–5 years. This constraint was reasonable, as Model 3 fit the data as well as Model 2b, $\Delta\chi^2(2) = 1.2$, $p = 0.55$.

Model 4 tested whether the covariances between the latent variable eta and the time-varying measures of EF could be constrained to 0, which would imply that the latent time-invariant influences on ADHD were independent of the effects of EF on ADHD. Model 4 fit the data more poorly than Model 3, $\Delta\chi^2(3) = 70.4$, $p < 0.001$. An inspection of the parameter estimates from Model 3 clearly indicated that all three of the covariances between the latent variable and time-varying measures of EF were non-zero, and no further constraints for the phi parameters were considered.

Continuing to build from Model 3 (because none of the constraints in Model 4 were supported), Model 5 tested whether the residual variances (epsilon) could be constrained to be equal, which would imply that the combined time-varying and time-invariant effects on ADHD were equivalent across ages 3–5 years. Model 5 fit the data more poorly than did Model 3, $\Delta\chi^2(2) = 31.1$, $p < 0.001$. An inspection of the parameter estimates from Model 3 suggested that the epsilon parameter at age 3 was larger than that at ages 4 and 5. Model

Table 2 Summary of model fit statistics and model comparisons

Model	Model fit statistics					Model comparisons	
	AIC	BIC	Chi-square (<i>df</i>)	RMSEA (90% CI)	CFI	Comparing models	Chi-square (<i>df</i>)
1	8019.89	8141.24	3.88 (3)	0.016 (0.000, 0.054)	0.999	–	–
2	8031.30	8142.54	13.78 (5)*	0.039 (0.015, 0.064)	0.992	1 vs. 2	8.57 (2)*
2b	8022.50	8138.79	6.98 (4)	0.025 (0.000, 0.056)	0.997	1 vs. 2b	2.76 (1)
3	8020.02	8126.20	8.25 (6)*	0.018 (0.000, 0.046)	0.998	2b vs. 3	1.19 (2)
4	8101.47	8192.48	75.60 (9)***	0.080 (0.064, 0.097)	0.938	3 vs. 4	70.38 (3)***
5	8063.45	8159.52	42.33 (8)***	0.061 (0.044, 0.079)	0.968	3 vs. 5	31.11 (2)***
5b	8027.85	8128.98	15.00 (7)*	0.031 (0.008, 0.053)	0.993	3 vs. 5b	5.70 (1)*

Model 3 provided the best fit to the data

AIC, Akaike information criteria; *BIC*, Bayesian information criteria; *CFI*, comparative fit index; *CI*, confidence interval; *RMSEA*, root mean squared error of approximation

* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$

5b freed the constraint on the epsilon parameter at age 3 but continued to fit the data more poorly than Model 3, $\Delta\chi^2(1) = 5.7, p = 0.017$.

Consistent with the nested likelihood ratio tests, a comparison of the information criteria across models also indicated that Model 3 represented the most parsimonious representation of the data (see Table 2). Hence, although all models met traditional standards for acceptable global model fit, the likelihood ratio tests and information criteria converged in identifying Model 3 as providing the best fit to the observed data. Having identified the model that best represented the observed data, we next focused on the interpretation of model parameters.

Effects of Executive Function Skills on ADHD Behaviors

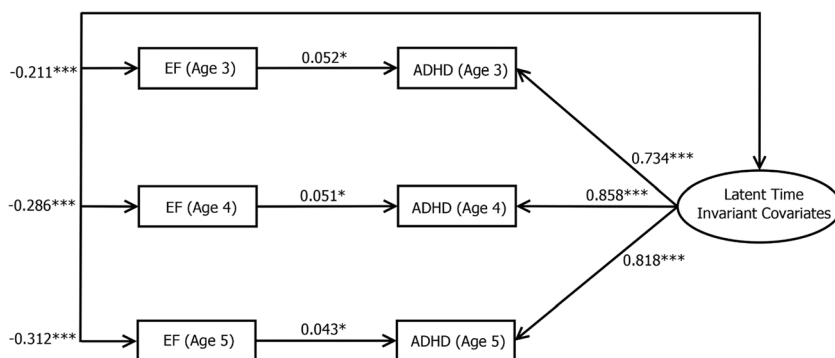
Based on Model 3, the within-person effect of EF on ADHD at age 3 was $\beta = -.052$ (95% CI = -0.093 to -0.011), $p = 0.012$; at age 4 was $\beta = -0.051$ (95% CI = -0.090 to -0.011), $p = 0.013$; and at age 5 was $\beta = -0.043$ (95% CI = -0.078 to -0.009), $p = 0.014$. In contrast, the between-person correlations involving the time-invariant latent influences on ADHD and time-specific indicators of EF were $\varphi_s = -0.211, -0.286,$ and -0.312 , for age 3, 4, and 5 years, respectively, all $p_s < 0.001$. Figure 2 replaces the general form of the panel model with model-based estimates. To reinforce the importance of identifying the model that provided the best fit to the observed data, we summarized the point estimate and confidence interval for the within-person effects of EF on ADHD across all the models in Fig. 3. It is noteworthy that the estimated effect of EF skills on ADHD behaviors was largest when the latent time-invariant influences on ADHD were assumed to be uncorrelated with time-varying measures of EF skills (Model 4), a condition that was not supported by the data. The ability to test this assumption is not possible in studies that employ cross-sectional designs or in studies that employ longitudinal designs but that do not parameterize models in ways that permit the explicit partitioning of between- and within-person variation in ADHD behaviors.

Supplementary Models

Up to this point, we have focused on the association between an overall measure of EF and ADHD. Although combining multiple EF tasks into a single score improves the reliability of measurement (Willoughby et al. 2017), ADHD may be more strongly associated with specific dimensions of EF skills, such as inhibitory control. To test this possibility, we recomputed the EF composite using up to three inhibitory control tasks and re-estimated the full sequence of structural equation models. The same pattern of results was evident with respect to the model comparisons. However, the magnitude of the effect was somewhat smaller for inhibitory control than for all EF tasks combined: age 3 $\beta = -0.037$ (95% CI = -0.077 to 0.003), $p = 0.068$; age 4 $\beta = -0.034$ (95% CI = -0.071 to 0.003), $p = 0.069$; age 5 $\beta = -0.028$ (95% CI = -0.059 to 0.002), $p = 0.071$. A summary of these results appears in supplementary materials (Online Resource Fig. 1 and Online Resource Table 1).

Up to this point, we have focused on overall ADHD behaviors. However, there is a long tradition in the ADHD literature of distinguishing inattentive from hyperactive-impulsive behaviors (Willcutt et al. 2012). We re-estimated the full sequence of structural equation models using EF composite scores with inattentive scores and then separately with hyperactive-impulsive scores. Model 3 continued to be the best representation of the data. None of the within-person associations between inattention and EF were statistically significant: age 3 $\beta = -0.036$ (95% CI = -0.078 to 0.007), $p = 0.100$; age 4 $\beta = -0.035$ (95% CI = -0.077 to 0.007), $p = 0.100$; age 5 $\beta = -0.030$ (95% CI = -0.067 to 0.006), $p = 0.103$. In contrast, the within-person associations between hyperactive-impulsive behaviors were small but statistically significant: age 3 $\beta = -0.064$ (95% CI = -0.106 to -0.023), $p = 0.003$; age 4 $\beta = -0.061$ (95% CI = -0.101 to -0.021), $p = 0.003$; age 5 $\beta = -0.053$ (95% CI = -0.088 to -0.018), $p = 0.003$. A summary of these results is available in supplementary materials (Inattentive: Online Resource Fig. 2 and Online Resource Table 2; Hyperactive/Impulsive: Online Resource Fig. 3 and Online Resource Table 3).

Fig. 2 Representation of Model 3 Estimates. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$



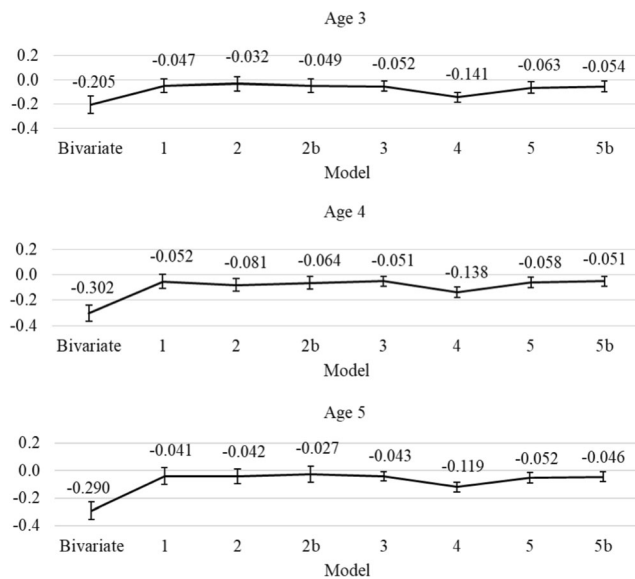


Fig. 3 Standardized Effects of EF Skills on ADHD Behaviors, by Model and Child Age

Discussion

A prevailing assumption in ADHD research has been that individual differences in ADHD symptomatology reflect the outward manifestation of underlying cognitive impairments. EF skills represent one domain of cognitive functioning that has been consistently related to ADHD symptoms across the life span. EF skills undergo pronounced normative improvements in early childhood, which has led to suggestions that early intervention efforts that target EF skill development may provide a means of preventing ADHD symptomatology or mitigating symptom-related impairments (Halperin et al. 2012; Sonuga-Barke and Halperin 2010). The purpose of this study was to leverage repeated-measures data to provide a stronger test of whether the association between EF skills and ADHD behaviors in early childhood reflects differences that exist between versus within children.

We observed bivariate associations between EF skills and ADHD behaviors that were in line with previous meta-analyses, including evidence of stronger associations among older versus younger preschool-aged children ($r_s = 0.30$ vs. 0.20). These age differences were consistent with the idea that the association between EF skills and ADHD is less pronounced during periods of substantial normative developmental change (Pauli-Pott and Becker 2015). Notably, we have previously established that our measures of EF skills and ADHD behaviors exhibited longitudinal measurement invariance during the ages studied here (Willoughby et al. 2012a, b, c). This helps to rule out the possibility that these age-related differences in the association between EF skills and ADHD behaviors is an artifact of poorer quality measurement at age 3 versus age 5 years.

Most studies that have investigated the association between EF skills and ADHD in early childhood have not provided

strong control for confounder variables that may account for some or all these associations. Here, we took advantage of repeated-measures data to control for all time-invariant confounder variables that might have accounted for the association between EF skills and ADHD behaviors. We fit a range of models that differed with respect to the assumptions that were made about the nature of the association between EF skills and ADHD behaviors, and we used formal statistical criteria to discern which model specification provided the best fit to the observed data. Our emphasis of model comparisons was important because all the models exhibited acceptable global model fit (see Table 2). Nonetheless, some models provided relatively better fit than others. We selected the model that provided the best representation of the observed data. It was only after we identified this model that we began to interpret parameter estimates. The benefits of this approach are evident from a review of Fig. 3, which demonstrates how the magnitude of the within-person association between EF and ADHD varies across a series of models, any one of which would have been considered to have “good enough fit” if considered in isolation.

In this study, the association between EF skills and ADHD behaviors primarily resulted from differences between children, which is consistent with confounding. That is, the results of this study suggest that EF skills are moderately to strongly correlated with other time-stable influences on children’s ADHD behaviors. When those time-stable influences were controlled for, the time-specific within-person time associations between EF skills and ADHD behaviors were small—far smaller than what is conveyed by meta-analytic studies (which typically conflate between- and within-person effects).

In many respects, the results of this study are unsurprising. The discordance between performance-based measures of EF skills and ADHD behaviors is well known, which is why performance-based EF measures are not used to diagnose ADHD. Paradoxically, a recurring idea in the literature, which has received some empirical support, is that efforts to improve a child’s EF skills (or related neurocognitive processes) will have corresponding benefits in terms of their ADHD behaviors (e.g., Spencer-Smith and Klingberg 2015). Although the methods of improving EF skills vary across studies—from meta-cognition training, to increased physical activity, to enhanced play, to adaptive computerized games—the theory of change underlying many interventions that are used in early childhood is the same (Halperin et al. 2013; Hoza et al. 2015; Tamm and Nakonezny 2015; van Dongen-Boomsma et al. 2014). This contradiction about whether an individual’s EF skills contribute to their ADHD behaviors—and if so, by how much—abounds in the literature. The results of this study suggest that most of the association between EF skills and ADHD behaviors derives from differences that exist between, not within, children.

It is instructive to contrast our approach with other recent studies that have investigated the association between ADHD

symptoms and EF skills in early childhood. For example, Sjöwall and colleagues reported that measures of EF skills that were obtained in early childhood predicted ADHD behaviors and academic achievement in adolescence, above and beyond the effects of preschool ADHD behaviors (Sjöwall et al. 2017). Pauli-Pott and colleagues reported that EF skills (and related neurocognitive processes) that were measured at ages 4 and 5 predicted ADHD symptom changes from age 4 to 5 but that the converse was not true (Pauli-Pott et al. 2017). Tibu and colleagues reported that psychosocial deprivation experienced in early childhood predicted EF skills and ADHD behaviors in middle childhood and adolescence and that EF skills served to mediate the association between early deprivation and later ADHD behaviors (Tibu et al. 2016a, 2016b). These studies and others like them are interpreted as evidence that individual differences in EF skills precede and contribute to the manifestation of ADHD behaviors, which is inherently a within-person inference. However, despite the use of longitudinal designs, none of these studies analyzed repeated-measures data in ways that explicitly distinguished between- and within-person sources of variation. As such, they offer a relatively weak basis of inference for the assertion that ADHD behaviors reflect underlying impairments in EF skills.

Another approach for investigating the association between EF skills and ADHD behaviors has involved recruiting samples of children who meet diagnostic criteria for ADHD and a matched group of typically developing comparison youth. All children complete a common set of EF (and/or other cognitive performance-based) tasks. Statistical methods are then used to identify subgroups of children who are homogenous with respect to their patterns of EF (and/or other cognitive) task performance (e.g., Fair et al. 2012; Rajendran et al. 2015; Roberts et al. 2017). These studies provide an elegant way of empirically documenting cognitive heterogeneity within ADHD youth, which was anticipated over a decade ago (see Nigg et al. 2005). Nonetheless, these studies do not inform questions about whether ADHD behaviors reflect underlying EF impairments, even for the subgroups of children who exhibit elevated ADHD symptomatology and who exhibit impairments in EF task performance. Once again, these studies do not test within-person associations.

This study suffered from five limitations. First, a variety of cognitive processes that have been implicated in the development of ADHD were not included in this study (e.g., delay aversion, reaction-time variability). It is unclear whether a similar pattern of results would have been evident had these other constructs been included. Second, we focused on the association between EF skills and ADHD behaviors, without consideration of whether these behaviors resulted in impairment, which may limit the clinical utility of this study. Third, our results indicated that stable inter-individual difference variables accounted for most of the association between EF skills and ADHD behaviors. However, we did not consider the

origin or predictors of these individual differences, which is an important direction for future research. Fourth, our study is based on a community sample of families and children who were recruited from nonmetropolitan, low-wealth counties in North Carolina and Pennsylvania. It is unclear to what extent these results would generalize to other settings. Fifth, we capitalized on normative changes that occur in EF skill development and ADHD behaviors in early childhood. The approach that we used here is not well suited to circumstances in which within-person variation in EF skills and ADHD behaviors are not expected.

Modern perspectives of developmental psychopathology emphasize the multilevel nature of all disorders, which represent complex associations among genetic, neural, cognitive, behavioral, and environmental factors. ADHD has long been characterized in this vein, such that individual differences in ADHD behaviors have been presumed to reflect developmental impairments in specific cognitive processes. Much of the extant evidence that has been marshalled to support this assumption conflates between- and within-person sources of variation, even though within-person variation is implicated in theoretical models of ADHD and is uniquely relevant to clinical decision making. We hope that the results of this study will challenge others to think critically about this distinction. Although we have emphasized the merits of using repeated-measures data, experimental studies that *narrowly* target improvements in EF (or other related cognitive) skills and that evaluate whether these experimentally induced improvements in EF skills result in improvements in ADHD behaviors, remain the strongest test of whether impaired EF skills contribute to the emergence and/or maintenance of ADHD behaviors in individual children.

Acknowledgments This study is part of the Family Life Project, which was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development grants P01 HD39667, with co-funding from the National Institute on Drug Abuse, and more recently from the National Institute of Health's Environmental influences on Child Health Outcomes Program (1UG3OD023332-01).

Compliance with Ethical Standards

Conflict of Interest None of the authors have conflicts of interest to report.

References

- Alderson, R. M., Rapport, M. D., & Kofler, M. J. (2007). Attention-deficit/hyperactivity disorder and behavioral inhibition: A meta-analytic review of the stop-signal paradigm. *Journal of Abnormal Child Psychology*, 35(5), 745–758. <https://doi.org/10.1007/s10802-007-9131-6>.
- Allison, P. D. (2009). *Fixed effects regression models*. Thousand Oaks, CA: Sage. <https://doi.org/10.4135/9781412993869>.

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders, 5th edition (DSM-V)*. Washington, DC: Author.
- Arias, V. B., Ponce, F. P., & Nunez, D. E. (2018). Bifactor models of attention-deficit/hyperactivity disorder (ADHD): An evaluation of three necessary but underused psychometric indexes. *Assessment, 25*(7), 885–897. <https://doi.org/10.1177/1073191116679260>.
- Arnett, A. B., Pennington, B. F., Young, J. F., & Hankin, B. L. (2016). Links between within-person fluctuations in hyperactivity/attention problems and subsequent conduct problems. *Journal of Child Psychology and Psychiatry, 57*(4), 502–509.
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*(1), 65–94.
- Bollen, K. A., & Brand, J. E. (2010). A general panel model with random and fixed effects: A structural equations approach. *Social Forces, 89*(1), 1–34. <https://doi.org/10.1353/sof.2010.0072>.
- Castellanos, F. X., & Aoki, Y. (2016). Intrinsic functional connectivity in attention-deficit/hyperactivity disorder: A science in development. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 1*(3), 253–261.
- Coghill, D. R., Seth, S., & Matthews, K. (2014). A comprehensive assessment of memory, delay aversion, timing, inhibition, decision making and variability in attention deficit hyperactivity disorder: Advancing beyond the three-pathway models. *Psychological Medicine, 44*(9), 1989–2001. <https://doi.org/10.1017/S0033291713002547>.
- Cortese, S. (2012). The neurobiology and genetics of attention-deficit/hyperactivity disorder (ADHD): What every clinician should know. *European Journal of Paediatric Neurology, 16*(5), 422–433. <https://doi.org/10.1016/j.ejpn.2012.01.009>.
- Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology, 62*(1), 583–619. <https://doi.org/10.1146/annurev.psych.093008.100356>.
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 63*(64), 135–168. <https://doi.org/10.1146/annurev-psych-113011-143750>.
- Doyle, A. E. (2006). Executive functions in attention-deficit/hyperactivity disorder. *The Journal of Clinical Psychiatry, 67*(Suppl 8), 21–26.
- Espy, K. (2017). The changing nature of executive control in preschool. *Monographs of the Society for Research in Child Development, 81*(4), 1–179. <https://doi.org/10.1111/mono.12273>.
- Fair, D. A., Bathula, D., Nikolas, M. A., & Nigg, J. T. (2012). Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. *Proceedings of the National Academy of Sciences of the United States of America, 109*(17), 6769–6774. <https://doi.org/10.1073/pnas.1115365109>.
- Garon, N., Bryson, S. E., & Smith, I. M. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134*(1), 31–60. <https://doi.org/10.1037/0033-2909.134.1.31>.
- Gomez, R., Vance, A., & Gomez, R. M. (2018). Validity of the ADHD Bifactor model in general community samples of adolescents and adults, and a clinic-referred sample of children and adolescents. *Journal of Attention Disorders, 22*(14), 1307–1319.
- Halaby, C. N. (2004). Panel models in sociological research: Theory into practice. *Annual Review of Sociology, 30*(1), 507–544. <https://doi.org/10.1146/annurev.soc.30.012703.110629>.
- Halperin, J. M., Bédard, A. C. V., & Curchack-Lichtin, J. T. (2012). Preventive interventions for ADHD: A neurodevelopmental perspective. *Neurotherapeutics: The Journal of the American Society for Experimental Neurotherapeutics, 9*(3), 531–541. <https://doi.org/10.1007/s13311-012-0123-z>.
- Halperin, J. M., Marks, D. J., Bedard, A. C., Chacko, A., Curchack, J. T., Yoon, C. A., & Healey, D. M. (2013). Training executive, attention, and motor skills: A proof-of-concept study in preschool children with ADHD. *Journal of Attention Disorders, 17*(8), 711–721. <https://doi.org/10.1177/1087054711435681>.
- Hervey, A. S., Epstein, J. N., & Curry, J. F. (2004). Neuropsychology of adults with attention-deficit/hyperactivity disorder: A meta-analytic review. *Neuropsychology, 18*(3), 485–503. <https://doi.org/10.1037/0894-4105.18.3.485>.
- Homack, S., & Riccio, C. A. (2004). A meta-analysis of the sensitivity and specificity of the Stroop color and word test with children. *Archives of Clinical Neuropsychology, 19*(6), 725–743. <https://doi.org/10.1016/j.acn.2003.09.003>.
- Hoza, B., Smith, A. L., Shoulberg, E. K., Linnea, K. S., Dorsch, T. E., Blazo, J. A., . . . McCabe, G. P. (2015). A randomized trial examining the effects of aerobic physical activity on attention-deficit/hyperactivity disorder symptoms in young children. *Journal of Abnormal Child Psychology, 43*(4), 655–667. <https://doi.org/10.1007/s10802-014-9929-y>.
- Huang-Pollock, C. L., Nigg, J. T., & Carr, T. H. (2005). Deficient attention is hard to find: Applying the perceptual load model of selective attention to attention deficit hyperactivity disorder subtypes. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 46*(11), 1211–1218. <https://doi.org/10.1111/j.1469-7610.2005.00410.x>.
- Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin, 144*(11), 1147–1185. <https://doi.org/10.1037/bul0000160>.
- Kofler, M. J., Irwin, L. N., Soto, E. F., Groves, N. B., Harmon, S. L., & Sarver, D. E. (2019). Executive functioning heterogeneity in pediatric ADHD. *Journal of Abnormal Child Psychology, 47*(2), 273–286. <https://doi.org/10.1007/s10802-018-0438-2>.
- Levy, F. (1991). The dopamine theory of attention deficit hyperactivity disorder (ADHD). *The Australian and New Zealand Journal of Psychiatry, 25*(2), 277–283. <https://doi.org/10.3109/00048679109077746>.
- Levy, F., Ban, C., & Sunohara, G. (1998). Directions of aetiological research on attention deficit hyperactivity disorder. *The Australian and New Zealand Journal of Psychiatry, 32*(1), 97–103. <https://doi.org/10.3109/00048679809062715>.
- Middeldorp, C. M., Hammerschlag, A. R., Ouwens, K. G., Groen-Blokhuis, M. M., St. Pourcain, B., Greven, C. U., et al. (2016). A genome-wide association meta-analysis of attention-deficit/hyperactivity disorder symptoms in population-based pediatric cohorts. *Journal of the American Academy of Child and Adolescent Psychiatry, 55*(10), 896–905. <https://doi.org/10.1016/j.jaac.2016.05.025>.
- Mostert, J. C., Onnink, A. M. H., Klein, M., Dammers, J., Harneit, A., Schulten, T., et al. (2015). Cognitive heterogeneity in adult attention deficit/hyperactivity disorder: A systematic analysis of neuropsychological measurements. *European Neuropsychopharmacology, 25*(11), 2062–2074. <https://doi.org/10.1016/j.euroneuro.2015.08.010>.
- Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide. Statistical analysis with latent variables (8th ed.)*. Los Angeles, CA: Muthén and Muthén. https://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf
- Nigg, J. T. (2001). Is ADHD a disinhibitory disorder. *Psychological Bulletin, 127*(5), 571–598.
- Nigg, J. T., Willcutt, E. G., Doyle, A. E., & Sonuga-Barke, E. J. (2005). Causal heterogeneity in attention-deficit/hyperactivity disorder: Do we need neuropsychologically impaired subtypes? *Biological Psychiatry, 57*(11), 1224–1230. <https://doi.org/10.1016/j.biopsych.2004.08.025>.
- Pauli-Pott, U., & Becker, K. (2011). Neuropsychological basic deficits in preschoolers at risk for ADHD: A meta-analysis. *Clinical*

- Psychology Review*, 31(4), 626–637. <https://doi.org/10.1016/j.cpr.2011.02.005>.
- Pauli-Pott, U., & Becker, K. (2015). Time windows matter in ADHD-related developing neuropsychological basic deficits: A comprehensive review and meta-regression analysis. *Neuroscience and Biobehavioral Reviews*, 55, 165–172. <https://doi.org/10.1016/j.neubiorev.2015.04.011>.
- Pauli-Pott, U., Schloß, S., Heinzel-Gutenbrunner, M., & Becker, K. (2017). Multiple causal pathways in attention-deficit/hyperactivity disorder – Do emerging executive and motivational deviations precede symptom development? *Child Neuropsychology*, 25, 1–19. <https://doi.org/10.1080/09297049.2017.1380177>.
- Pelham, W. E., et al. (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(2), 285–299.
- Rajendran, K., O'Neill, S., Marks, D. J., & Halperin, J. M. (2015). Latent profile analysis of neuropsychological measures to determine preschoolers' risk for ADHD. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 56(9), 958–965. <https://doi.org/10.1111/jcpp.12434>.
- Roberts, B. A., Martel, M. M., & Nigg, J. T. (2017). Are there executive dysfunction subtypes within ADHD? *Journal of Attention Disorders*, 21(4), 284–293. <https://doi.org/10.1177/1087054713510349>.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147–177. <https://doi.org/10.1037/1082-989x.7.2.147>.
- Schoemaker, K., Mulder, H., Deković, M., & Matthys, W. (2013). Executive functions in preschool children with externalizing behavior problems: A meta-analysis. *Journal of Abnormal Child Psychology*, 41(3), 457–471. <https://doi.org/10.1007/s10802-012-9684-x>.
- Sjöwall, D., Bohlin, G., Rydell, A. M., & Thorell, L. B. (2017). Neuropsychological deficits in preschool as predictors of ADHD symptoms and academic achievement in late adolescence. *Child Neuropsychology*, 23(1), 111–128. <https://doi.org/10.1080/09297049.2015.1063595>.
- Sjöwall, D., & Thorell, L. B. (2019). A critical appraisal of the role of neuropsychological deficits in preschool ADHD. *Child Neuropsychology*, 25(1), 60–80. <https://doi.org/10.1080/09297049.2018.1447096>.
- Sonuga-Barke, E., Bitsakou, P., & Thompson, M. (2010). Beyond the dual pathway model: Evidence for the dissociation of timing, inhibitory, and delay-related impairments in attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(4), 345–355. <https://doi.org/10.1097/00004583-201004000-00009>.
- 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, and Allied Disciplines*, 51(4), 368–389. <https://doi.org/10.1111/j.1469-7610.2009.02195.x>.
- Spencer-Smith, M., & Klingberg, T. (2015). Benefits of a working memory training program for inattention in daily life: A systematic review and meta-analysis. *PLoS One*, 10(3), e0119522. <https://doi.org/10.1371/journal.pone.0119522>.
- Swanson, J., Castellanos, F. X., Murias, M., LaHoste, G., & Kennedy, J. (1998). Cognitive neuroscience of attention deficit hyperactivity disorder and hyperkinetic disorder. *Current Opinion in Neurobiology*, 8(2), 263–271. [https://doi.org/10.1016/S0959-4388\(98\)80150-5](https://doi.org/10.1016/S0959-4388(98)80150-5).
- Tamm, L., & Nakonezny, P. A. (2015). Metacognitive executive function training for young children with ADHD: A proof-of-concept study. *Attention Deficit and Hyperactivity Disorders*, 7(3), 183–190. <https://doi.org/10.1007/s12402-014-0162-x>.
- Tamm, L., Narad, M. E., Antonini, T. N., O'Brien, K. M., Hawk, L. W., Jr., & Epstein, J. N. (2012). Reaction time variability in ADHD: A review. *Neurotherapeutics: The Journal of the American Society for Experimental Neurotherapeutics*, 9(3), 500–508. <https://doi.org/10.1007/s13311-012-0138-5>.
- Tibu, F., Sheridan, M. A., McLaughlin, K. A., Nelson, C. A., Fox, N. A., & Zeanah, C. H. (2016a). Disruptions of working memory and inhibition mediate the association between exposure to institutionalization and symptoms of attention deficit hyperactivity disorder. *Psychological Medicine*, 46(3), 529–541. <https://doi.org/10.1017/S0033291715002020>.
- Tibu, F., Sheridan, M. A., McLaughlin, K. A., Nelson, C. A., Fox, N. A., & Zeanah, C. H. (2016b). Reduced working memory mediates the link between early institutional rearing and symptoms of ADHD at 12 years. *Frontiers in Psychology*, 7(1850). <https://doi.org/10.3389/fpsyg.2016.01850>.
- van Dongen-Boomsma, M., Vollebregt, M. A., Buitelaar, J. K., & Slaats-Willemse, D. (2014). Working memory training in young children with ADHD: A randomized placebo-controlled trial. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 55(8), 886–896. <https://doi.org/10.1111/jcpp.12218>.
- 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(5), 1–150, vii. <https://doi.org/10.1111/mono.12046>.
- Wahlstedt, C., Thorell, L. B., & Bohlin, G. (2009). Heterogeneity in ADHD: Neuropsychological pathways, comorbidity and symptom domains. *Journal of Abnormal Child Psychology*, 37(4), 551–564.
- Walton, E., Pingault, J. B., Cecil, C. A., Gaunt, T. R., Relton, C. L., Mill, J., & Barker, E. D. (2017). Epigenetic profiling of ADHD symptoms trajectories: A prospective, methylome-wide study. *Molecular Psychiatry*, 22(2), 250–256. <https://doi.org/10.1038/mp.2016.85>.
- Wender, P. H. (1975). The minimal brain dysfunction syndrome. *Annual Review of Medicine*, 26(1), 45–62. <https://doi.org/10.1146/annurev.me.26.020175.000401>.
- Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological Psychiatry*, 57(11), 1336–1346. <https://doi.org/10.1016/j.biopsych.2005.02.006>.
- Willcutt, E. G., Nigg, J. T., Pennington, B. F., Solanto, M. V., Rohde, L. A., Tannock, R., . . . Lahey, B. B. (2012). Validity of DSM-IV attention deficit/hyperactivity disorder symptom dimensions and subtypes. *Journal of Abnormal Psychology*, 121(4), 991–1010. doi:<https://doi.org/10.1037/a0027347>.
- 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(2), 306–317. <https://doi.org/10.1037/a0018708>.
- Willoughby, M., & Blair, C. (2011). Test-retest reliability of a new executive function battery for use in early childhood. *Child Neuropsychology*, 17(6), 564–579. <https://doi.org/10.1080/09297049.2011.554390>.
- Willoughby, M. T., Blair, C. B., Wirth, R. J., Greenberg, M., & the Family Life Project Investigators. (2012a). The measurement of executive function at age 5: Psychometric properties and relationship to academic achievement. *Psychological Assessment*, 24(1), 226–239. <https://doi.org/10.1037/a0025361>.
- Willoughby, M. T., Pek, J., Greenberg, M. T., & the Family Life Project Investigators. (2012b). Parent-reported attention deficit/hyperactivity symptomatology in preschool-aged children: Factor structure, developmental change, and early risk factors. *Journal of Abnormal Child Psychology*, 40(8), 1301–1312. <https://doi.org/10.1007/s10802-012-9641-8>.

- Willoughby, M. T., Wirth, R. J., Blair, C. B., & the Family Life Project Investigators. (2012c). Executive function in early childhood: Longitudinal measurement invariance and developmental change. *Psychological Assessment, 24*(2), 418–431. <https://doi.org/10.1037/a0025779>.
- Willoughby, M. T., & Blair, C. B. (2016). Longitudinal measurement of executive function in preschoolers. In J. Griffin, L. Freund, & P. McCardle (Eds.), *Executive function in preschool age children: Integrating measurement, neurodevelopment and translational research* (pp. 91–113). Washington, DC: American Psychological Association Press. <https://doi.org/10.1037/14797-005>.
- Willoughby, M. T., Blair, C. B., & the Family Life Project Investigators. (2016). Measuring executive function in early childhood: A case for formative measurement. *Psychological Assessment, 28*(3), 319–330. <https://doi.org/10.1037/pas0000152>.
- Willoughby, M. T., Kuhn, L. J., Blair, C. B., Samek, A., & List, J. A. (2017). The test-retest reliability of the latent construct of executive function depends on whether tasks are represented as formative or reflective indicators. *Child Neuropsychology, 23*(7), 822–837.
- Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge: MIT Press.
- Zametkin, A. J., & Rapoport, J. L. (1987). Neurobiology of attention deficit disorder with hyperactivity: Where have we come in 50 years? *Journal of the American Academy of Child and Adolescent Psychiatry, 26*(5), 676–686. <https://doi.org/10.1097/00004583-198709000-00011>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.