A Cross-Lagged Model of the Development of ADHD Inattention Symptoms and Rapid Naming Speed

Anne B. Arnett · Bruce F. Pennington · Erik Willcutt ·
Julia Dmitrieva · Brian Byrne · Stefan Samuelsson ·
Richard K. Olson

Abstract Although previous research has identified contemporaneous associations between cognitive deficits and symptom phenotypes in Attention Deficit/Hyperactivity Disorder, no studies have as yet attempted to identify direction of effect. The present study used cross-lagged path modeling to examine competing hypotheses about longitudinal associations between rapid naming speed and symptoms of inattention in children. 1,506 school-age twins from Australia and the U.S. were tested for inattention, hyperactivity/impulsivity, and rapid naming speed at three and four time points, respectively. Symptom severity of inattention from Kindergarten to fourth grade is consistently predicted by previous rapid naming, over and above auto-regressive and correlational associations in the model. Likewise, inattention symptoms have a small but significant predictive effect on subsequent rapid naming. The findings support a reciprocal relationship between naming speed and ADHD inattentive symptoms.

Keywords ADHD · Rapid naming speed · Processing speed · Structural equation modeling · Longitudinal

A. B. Arnett (✉) · B. F. Pennington · J. Dmitrieva
Department of Psychology, University of Denver,
Frontier Hall. 2155 S. Race St.,
Denver, CO 80208, USA
e-mail: awbernard@gmail.com

E. Willcutt · R. K. Olson
Department of Psychology, University of Colorado,
Boulder, CO, USA

B. Byrne
Department of Psychology, University of New England,
Armidale, Australia

B. Byrne · S. Samuelsson · R. K. Olson
Department of Psychology, University of Linkoping,
Linkoping, Sweden

S. Samuelsson
Department of Psychology, Stavanger University,
Stavanger, Norway

Current literature on the neuropsychology of ADHD highlights a need to identify cognitive endophenotypes of this disorder (Castellanos and Tannock 2002). While several candidate endophenotypes have been identified for ADHD (Bidwell, Willcutt, DeFries, and Pennington 2007a), a key test of an endophenotype has rarely been performed, namely that it precedes and predicts later ADHD. Although we tend to assume that cognitive deficits associated with disorders exhibit this relation, other relations are clearly possible, as discussed below. Therefore, candidate endophenotypes need to be tested longitudinally. For the purposes of this study, an endophenotype is defined as a cognitive precursor that 1) explains some variance in symptoms of a disorder, and 2) is genetically correlated with the disorder.

Hence, the overall goal of this study was to examine the direction of association between individual differences in ADHD inattention symptoms and rapid automatized naming speed (RAN), a measure shown to be deficient in children with ADHD (e.g. Nigg, Blaskey, Huang-Pollock, and

Rappley 2002; Tannock, Martinussen, and Frijters 2000) and highly correlated with processing speed (Chhabildas, Pennington, and Willcutt 2001; McGrath et al. 2011), another known cognitive correlate of ADHD. In order to test the direction of effects between RAN and inattention symptoms, we modeled longitudinal relationships from pre-Kindergarten through fourth grade using an autoregressive, cross-lagged path model design. This type of model allows for simultaneous examination of longitudinal influences of one construct on another, and vice versa, while also controlling for contemporaneous associations between constructs, and the stability of each construct over time. The current study tests four competing hypotheses regarding the direction of effects between RAN and inattention symptoms: 1) RAN and inattention symptoms demonstrate reciprocal, longitudinal effects; 2) RAN has a direct, longitudinal effect on inattention symptoms; 3) inattention symptoms have a direct effect on RAN; or 4) inattention symptoms and RAN do not influence each other directly, but share variance caused by a third, unmeasured factor.

Attention Deficit/Hyperactivity Disorder

Attention Deficit/Hyperactivity Disorder (ADHD) is one of the most common neurodevelopmental disorders, affecting 3 to 5% of school-age children in a 6-month period (Pennington 2002; Satcher 2000). This disorder is divided into three subtypes based on differential elevations of inattention and hyperactivity-impulsivity symptoms: inattentive type (ADHD-I), hyperactive-impulsive type (ADHD-HI), and combined type (ADHD-C) (American Psychiatric Association 2000). The distinction between inattention and hyperactivity/impulsivity symptoms is strongly supported by factor analytic studies and differential associations with social, academic, and occupational impairment and concurrent mental disorders (see review by Willcutt et al. 2012). Further, previous neuropsychological studies consistently found that attention deficits are more strongly associated with weaknesses in a range of neuropsychological domains than are hyperactivity/impulsivity symptoms. Therefore, the current study hypothesizes a unique longitudinal association between RAN and ADHD inattentive symptoms, specifically.

ADHD is widely considered to be an extreme phenotype, meaning that symptoms of inattention and hyperactivity/impulsivity in these individuals are at the extreme end of a continuum (Pennington 2002). It follows that the etiology of symptoms in the normal range would not differ from those symptoms in either the extremely high- or low-functioning ranges. Consequently, tests of the relation between a candidate endophenotype and a symptom dimension

can be done by examining the relation between individual differences in rapid naming and inattention symptoms in a population sample. Previous research in support of the extreme phenotype theory has found that rapid naming predicts behavior at both the adaptive (attentive) and maladaptive (inattentive) ends of the ADHD symptom spectrum in the current sample (Arnett et al. 2011).

Naming and Processing Speed as Cognitive Correlates of ADHD

Previous research has highlighted the importance of accurately identifying a distinct neuropsychological profile of childhood ADHD (Bidwell, Willcutt, DeFries, and Pennington 2007b; Chhabildas et al. 2001; Nigg et al. 2002). Greater understanding of specific cognitive deficits would lead to better evaluation and identification, improved treatment, and more informed investigations of etiology. Further, knowledge about the developmental trajectory of these cognitive correlates in relation to the behavioral deficits has potential to improve targeted, developmentally appropriate evaluation and treatment across different ages.

The present study aims to describe the longitudinal associations between inattention symptoms and one of the cognitive correlates of ADHD, rapid naming speed. Deficits in rapid naming speed and nonverbal processing speed are two of the most consistent findings in ADHD (e.g. Kalff et al. 2005; Shanahan et al. 2006; Weiler, Bernstein, Bellinger, and Waber 2000; Willcutt et al. 2007). The association between ADHD symptoms and slow cognitive speed remains significant even when measures of other replicated deficits such as response inhibition, response variability, and working memory are also included in the model. Therefore, although slow cognitive speed does not appear to be necessary or sufficient to cause ADHD (e.g. Nigg 2005; Sonuga-Barke 2005; Willcutt, Pennington, Olson, Chhabildas, and Hulslander 2005), it is an important part of a comprehensive neuropsychological model of ADHD. Although measures of nonverbal processing speed (e.g. WISC-R Coding) and measures of verbal naming speed (e.g. Rapid Automatized Naming) load on separate factors in some samples, these constructs are highly correlated (e.g., $r=0.77$; McGrath et al. 2011). Further, rapid naming speed tasks have been shown to be an important component of discriminant function analysis batteries for distinguishing ADHD and non-ADHD groups (Carte, Nigg, and Hinshaw 1996; Tannock et al. 2000). Thus, a better understanding of how rapid naming skills develop in relation to inattentive behavior is both clinically and scientifically relevant.

Previous Longitudinal Studies of ADHD

Test-retest reliability of ADHD symptoms is high for both inattention and hyperactivity/impulsivity dimensions over periods less than 1 year (see Willcutt et al. 2012) and only slightly lower over longer intervals ranging from 1 to 5 years (r=0.64 for both dimensions), suggesting that the rank ordering of individuals in the population on both symptom dimensions is largely stable over time. Although the overall position of each individual in the population distribution appears to be relatively stable, longitudinal studies indicate that inattention and hyperactivity/impulsivity symptoms may follow somewhat different developmental trajectories. For example, one prospective longitudinal study finds that children first diagnosed with DSM-IV ADHD in preschool exhibited a significant age-related decline in hyperactive/impulsive behaviors over the first nine years of the study. This decline was not related to pharmacologic or psychosocial treatment (Lahey et al. 1998; Lahey et al. 2004; Lahey, Pelham, Loney, Lee, and Willcutt 2005; Lahey and Willcutt 2010). In contrast, symptom levels of inattention did not change significantly over the course of the study. Similar results were reported in a five-year follow-up study of a sample of females with ADHD who were first assessed between 6 and 12 years of age (Hinshaw, Owens, Sami, and Fargeon 2006) and in a population-based longitudinal study that followed children from 8 to 17 years of age (Larsson, Lichtenstein, and Larsson 2006).

Despite ample research on the trajectory of ADHD symptoms, longitudinal studies of the cognitive correlates of ADHD are scarce. Wahlstedt, Thorell and Bohlin (2008) examined the associations between preschoolers' executive functioning (EF) skills (inhibition, working memory, and verbal fluency) and ADHD symptoms 2 years later. They found that as a group, children with EF deficits had worse inattention symptoms 2 years later, but there was no effect in the opposite direction. However, the effect sizes of the first finding were small, and significances were reduced when controlling for IQ. Likewise, Berlin and colleagues (2003) found that inhibitory control in male preschoolers was related to later inattention, working memory, regulation, and reconstitution at school age. This study also found significant initial concurrent relationships between cognitive correlates and inattention in boys; however, the authors did not test for the opposite direction of effects, nor did they control for autoregressive associations within each construct. Brocki and Bohlin (2006) conducted a cross-sectional study to test a developmental theory of ADHD and cognitive functions associated with executive functioning. They found that a factor of verbal working memory (digit span forwards and backwards) was significantly related to inattention in a group of 9.8–13 year olds, but not in the 6–9.7 year old age group. Trends toward a

similar interaction effect were found for non-verbal working memory and speed/arousal (reaction times) factors. In contrast, the younger age group exhibited an age-specific deficit in inhibition. A second study by Brocki and colleagues (2007) failed to find any relationship, either longitudinal or concurrent, between ADHD symptoms and working memory in preschoolers who were followed up 2 years later.

These data are consistent with the prominent theory that identify executive functions, working memory, and reaction time as candidate endophenotypes of ADHD (e.g., Barkley 1997), but do not clarify the developmental association between inattention and cognitive deficits during childhood. Overall, previous longitudinal and cross-sectional studies suggest that the relation between ADHD symptoms and cognitive functioning may change in important ways over childhood and adolescence, but the inconsistent pattern of results suggests that additional research is needed to clarify these associations. Further, previous studies have largely focused on executive functions and have not tested associations with cognitive speed.

As the literature review reveals, there are few longitudinal studies of cognitive predictors of ADHD. Further, none that we know of have employed a cross-lagged design or path modeling to test the direction of effects that is generally accepted in the literature, namely that earlier cognitive dysfunction predicts later inattention symptoms rather than the other way around.

Methods

Participants

Subjects were recruited as part of the International Longitudinal Twin Study (ILTS) and included 1,506 twins from the National Health and Medical Research Council's Australian Twin Registry (N=528) and from the Colorado Twin Registry in the United States (N=978). This sample has been described in detail elsewhere (see Samuelsson et al. 2005; Willcutt et al. 2007) so the following description will be brief. At Time 1, all subjects were in their final pre-Kindergarten (Pre-K) year, with ages ranging from 47 to 70 months (M=57.3) in Australia and 54 to 71 months (M=58.8) in the United States. Time 2 testing was conducted near the end of Kindergarten (K). Time 3 measures were administered near the end of 1st grade. RAN measures were not collected in 2nd grade, and no testing was done at the end of 3rd grade, so these time points were not included in path analyses. However, reading skills were measured at 2nd grade for classification of subjects at risk for a reading disorder. This study is ongoing, and additional data were also available at the end of 4th grade for the U.S. sample only. Overall, the sample

included 749 (49 %) females, 656 non-white subjects (Asian, Black, Hispanic, or Pacific Islander), 530 subjects who did not report ethnicity, and a mean parental education of approximately 14 years (see Table 1).

Procedure

Procedures for the International Longitudinal Twin Study have been described in detail elsewhere (Samuelsson et al. 2005; Willcutt et al. 2007), so the following description will focus on the assessments relevant to this article. Testing was performed at home or at school. In addition to the measures of rapid naming speed and ADHD symptoms considered in this study, subjects were tested for various skills related to reading proficiency. Tests were administered individually to each child over the course of five, one-hour sessions at the pre-K time point, and during single one-hour sessions thereafter. Each twin in a pair was tested at the same time by different testers. Parent ratings of ADHD symptoms were obtained via mail in Australia, and in person while individual testing was completed in the U.S.

Measures

Behavioral Measures Inattention and hyperactivity/impulsivity symptom factor scores were calculated using established measures of ADHD at each time point. The Disruptive Behavior Rating Scale (DBRS; Barkley and Murphy 1998) was used to obtain parent and tester ratings of the 18 symptoms of DSM-IV ADHD at Pre-K, K, 1st and 4th grades. Each symptom on the DBRS is rated on a four-point scale: 0=*never or rarely,* 1=*sometimes,* 2=*often,* and 3=*very often*). Inattention, hyperactive/impulsivity, and total ADHD composite scores were created by summing the individual item scores within each dimension. Previous results from these samples indicate that ratings on the DBRS are internally consistent (Willcutt et al. 2007). At pre-K, DBRS ratings were collected from parents once, and tester ratings were collected at each of the five assessment sessions. At K, 1st and 4th grades, parent and tester each completed the DBRS once. At K-4th grades ADHD factor scores also included parent ratings of the Strengths and Weaknesses of ADHD Symptoms and Normal Behavior (SWAN). The SWAN also targets DSM-IV criterion for ADHD symptoms, but is scored on a broader scale that allows for variance at the high (non-symptomatic) ends of inattention and hyperactivity/impulsivity behaviors as well as the symptomatic ends (Arnett et al. 2011; Swanson et al. 2006). As with the DBRS, item scores were summed within dimensions to obtain raw scale scores prior to factor analysis. Cronbach's alpha coefficient reliability for inattention scales (for all available measures and reporters) at pre-K, K and 1st grade ranged from $\alpha=0.618$–0.844; hyperactivity/impulsivity ranged from $\alpha=0.566$–0.830. Alpha coefficients for the 4th grade time point, which only included U.S. data, were significantly lower: inattention $\alpha=0.088$; hyperactivity/impulsivity $\alpha=0.079$.

ADHD diagnostic status was determined by calculating the means of inattention and hyperactivity/impulsivity factor scores at all available time points. Subjects who were in the highest 10th percentile for inattention or hyperactivity/impulsivity factor scores were labeled as ADHD-I or ADHD-HI, respectively; subjects at the 10th percentile for both subscales were labeled as ADHD-C. ADHD raw symptom severity (used for group comparisons,

Table 1 Sample characteristics

	ADHD-I	ADHD-HI	ADHD-C	None	Total
N	84	84	66	1272	1506
% Female	44 %	37 %	20 %	52 %	49 %
% Australia	33 %	21 %	20 %	37 %	35 %
% MZ twins	41 %	49 %	47 %	51 %	50 %
% Non-White	44 %	27 %	21 %	46 %	44 %
Birthweight in grams *M*(SD)	2321 (581)	2427 (518)	2255 (577)	2401 (524)	2390 (530)
Parental education in years *M*(SD)	14.03 (2.10)	13.92 (2.21)	13.52 (2.37)	13.99 (2.11)	13.97 (2.13)
Inattention Score, 2nd Grade DBRS *M*(SD)	11.03 (4.53)	7.41 (3.92)	13.44 (4.99)	4.18 (3.58)	5.32 (4.60)
Hyperactivity Score, 2nd Grade DBRS *M*(SD)	4.98 (3.57)	10.14 (4.47)	11.47 (5.21)	3.32 (3.34)	4.32 (4.31)
ADHD Total Score, 2nd Grade DBRS *M*(SD)	16.02 (7.16)	17.56 (7.41)	24.90 (8.91)	7.50 (5.91)	9.63 (7.94)
TOWRE Standard Score *M*(SD)	93.35 (13.89)	103.03 (14.07)	92.31 (12.70)	105.67 (13.90)	104.22 (14.38)

ADHD-I = Highest 10th percentile for mean Inattention factor scores; ADHD-HI = Highest 10th percentile for mean Hyperactivity/Impulsivity factor scores; ADHD-C = In highest 10th percentile for both mean Inattention and mean Hyperactivity/Impulsivity factor scores. None = below 10th percentile for all subscales. Non-White = Asian, Black, Hispanic, Pacific Islander. ADHD subscale and total scores are raw scores derived from DBRS parent ratings in 2nd grade. Mean TOWRE Standard Scores are derived from averaging standard scores on forms A and B for Sight Word Efficiency and Phonemic Decoding subtests

see Table 1) was calculated by totaling item ratings within each scale on the parent-report DBRS at 2nd grade. 2nd grade was the chosen time point for this analysis because it was the latest time point with both U.S. and Australian data available.

Given the strong association between inattention and hyperactivity/impulsivity, in our path analyses we controlled for one behavioral construct in the other by regressing the factor score for the target construct onto that of the non-target construct. The residual scores were used in all subsequent analyses. By isolating the variance contributing to inattention symptoms without hyperactivity/impulsivity symptoms, and vice versa, we were able to test and contrast the unique longitudinal relationships between RAN and each of these behavioral factors.

Cognitive Measures Rapid naming factor scores were derived using measures of rapid automatized naming at each time point. The Rapid Naming subtests from the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, and Rashotte 1999) were used as measures of RAN. The specific subtests administered at each time point were chosen according to developmental level (Wagner et al. 1999). Thus, at Pre-K, subjects completed Rapid Object Naming and Rapid Color Naming. At K, subjects completed two Rapid Color Naming, two Rapid Digit Naming, and two Rapid Letter Naming tests. At 1st and 4th grades, subjects completed two Rapid Digit Naming and two Rapid Letter Naming subtests. In each subtest, the subject was asked to name items on a page, out loud, as quickly as possible. Scores were recorded as time (in seconds) to complete the page. RAN factor scores were derived from these raw time scores.

Given that there is a phoneme awareness component to the CTOPP tests (Vukovic and Siegel 2006; Wagner et al. 1999), we performed additional, comparative analyses in which we regressed a phoneme awareness factor out of the RAN score to create a more specific measure of response speed. Phoneme awareness factors were comprised of one elision and one blending test at each time point from Pre-K to 1st grade. Elision tests require the child to delete a syllable or phoneme from a word to form a new word; blending tests require the child to combine phonemes or syllables to form a familiar word. In pre-K, both tests were a pre-reading measure designed for this study and previously described by Byrne, et al. (2002). In Kindergarten and 1st grades, the elision and blending tests came from the CTOPP. Scores were recorded as the number correct and factor scores were derived from these raw scores.

In order to test whether the models fit well regardless of reading skill, a reading score was calculated by averaging standardized scores for Sight Word Efficiency and Phonemic Decoding subtests, versions A and B, from the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, and Rashotte 1999) at 2nd grade. Second grade scores were used because this was the oldest time point where subjects from both countries had reading data. TOWRE scores are standardized with a mean =100, standard deviation =15. Subjects who scored in the lowest 10th percentile of the sample were categorized as being at risk for a reading disorder. Mean TOWRE standardize score for the reading disorder risk group was 78.48, SD=5.66; Mean TOWRE score for the non-risk group was 107.09, SD=12.00.

Data Cleaning

Transformations DBRS scale scores were noted to have non-normal distributions, with skew values ranging from 0.94 to 1.62, and kurtosis values from 0.78 to 3.0. Thus, consistent with the literature, these data were log transformed to approximate a normal distribution before proceeding with factor analyses (Hartman, Rhee, Willcutt, and Pennington 2007). The distributions of the transformed variables were in the acceptable range (skew: -1.04 to 0.07, kurtosis: -1.04 to 1.45).

Missing Data Mplus 6.0 uses full information maximum likelihood (FIML) to account for missing data. At each time point, the number of subjects with missing data on one or more variables was: Pre-K: 16 (1 %), Kindergarten: 226 (15 %), 1st grade: 297 (20 %), 4th grade for U.S. sample only: 104 (11 %). Subjects with missing data at any time point had a mean parental education 0.44 years lower on average ($t=3.83$, $p<0.001$) and a birth weight 0.16 pounds lower on average ($t=2.043$, $p<0.05$). Subjects with any missing data were less likely to fall into the top 10th percentile for any of the ADHD subtypes ($t=3.302$, $p<0.01$). There were no group differences in gender or race for subjects with missing data.

Factor Scores As described above, ADHD symptoms were assessed using different methods at baseline versus the follow-up assessments. At the pre-K time point, ADHD symptoms were measured five times by testers and once by parents, whereas the kindergarten-through-4th grade interviews included a tester report using the DBRS, a parent report using the DBRS, and an additional parent report using the SWAN. In order to study these constructs over time, we employed two-level confirmatory analyses, using maximum likelihood estimation with robust standard errors, to create single inattention and hyperactivity factor scores for each time point. Two-level models were used in order to account for non-independence of the twin data. The confirmatory factor analysis model included a single factor for

each construct (i.e., attention or hyperactivity/impulsivity) with six indicators (tester DBRS at each assessment plus one parent rating) for pre-K data and three indicators (tester DBRS, parent DRBS, and parent SWAN) for kindergarten through 4th grade data. The resulting factor scores were saved and used for subsequent analyses. Similarly, RAN and Phoneme Awareness were assessed with various subtests that changed as children grew older and required age-appropriate test content. Thus, subtests administered at each time point (described above) were combined into a single RAN or phoneme awareness score for that time point using two-level confirmatory factor analyses, and the resulting factor scores were used in subsequent analyses.

Power Power estimates for path models are not well defined in the literature. Methodological guidelines provided by MacCallum et al. (1996) suggest that our most complex model would require 500 subjects to achieve 80 % power to detect an exact model fit (defined as RMSEA between .00 and 0.05). Alternative guidelines by Kline (2005) suggest that only 320 subjects are needed. The current analyses included 1,506 subjects (753 independent subjects). Thus, by either guideline, our sample size was estimated to provide an acceptable level of power. Given the large sample size and number of models tested, we considered path coefficients with p values less than 0.01 to be statistically significant.

Data Analysis

This study utilized a cross-lagged, structural equation modeling design to test the fit of four alternative models. Two-level path models utilized factor scores rather than latent variables and accounted for non-independence of observations at the twin level. The models were analyzed with Mplus 6.0. Mplus was not capable of simultaneously modeling a longitudinal cross-lagged model that both used latent variables and accounted for nesting within twins. Thus, as described above, we computed factor scores from the two-level confirmatory analysis model and used these factor scores, rather than latent variables, to model the family relations in a two-level cross-lagged model.

Path modeling is well suited for studying neuropsychological development, due to its capacity to simultaneously estimate direct and indirect associations in longitudinal data (Hays, Marshall, Wang, and Sherbourne 1994). The advantage of path modeling over another multivariate analysis, such as multiple regression, is that multiple factors (i.e. ADHD symptoms and rapid naming speed) can be entered into the equation simultaneously, as opposed to being forced into a temporal sequence. The autoregressive, cross-lagged panel model estimates the association between RAN and attention over time. Autoregressive path weights account for the stability of each measure across two consecutive time points (i.e. paths a' and b' in Fig. 1), while the contemporaneous correlations between the two factors (c' and d' in Fig. 1) are also estimated.

Path models of longitudinal data were used to test four competing models of the developmental association between rapid automatized naming speed and inattention symptoms. The first model is a reciprocal model that suggests that RAN and inattention symptoms have bi-directional associations, with early inattention symptoms predicting later RAN and early RAN predicting later inattention symptoms. The second model proposes that individual differences in early RAN predict later inattention symptoms, whereas early inattention symptoms do not predict later RAN. In contrast, the third model suggests that individual differences in inattention symptoms predict later RAN, but RAN does not predict later inattention symptoms. Finally, in the fourth model there is a significant concurrent association between inattention symptoms and RAN, but not the predictive associations from one construct to the other at a later time point.

The cross-lag paths that are included in the model (e' and f' in Fig. 1) indicate the extent to which early RAN or inattention (INATT) symptom factor scores predict scores on the other measure at a later time point, independent of the longitudinal correlations between measures of the same construct and the contemporaneous correlation between the constructs at each time point. The fit of the four competing models is tested by dropping different sets of the cross-lag paths and calculating the chi-square differences. The reciprocal model (#1) includes all cross-lag paths, implying a bi-directional relationship between the factors. Models 2–4 are then tested by removing specific cross-lag paths from the saturated reciprocal model. In model #2 path e' is dropped because this model predicts that early RAN will predict later INATT but early INATT will not predict later RAN, whereas path f' is dropped in model #3 because it makes the opposite prediction. Finally, all cross-lag paths are dropped in model #4 because this model suggests that the relation between RAN and INATT is fully explained by a third factor that is not included in the model.

Chi square difference testing was utilized to compare the model fits. Due to the number of models tested, a chi square change that was significant at the $p<0.01$ level was considered significant. Per the recommendations of the authors of Mplus, the Chi Square difference was calculated using the Satorra-Bentler Scaled Chi Square. This value utilizes a correction factor to account for non-normality (Satorra 2000). If two models fit the data equally well, the most parsimonious model was chosen, consistent with literature on path modeling (Kline 2005).

Fig. 1 Autoregressive, cross-lagged panel path model
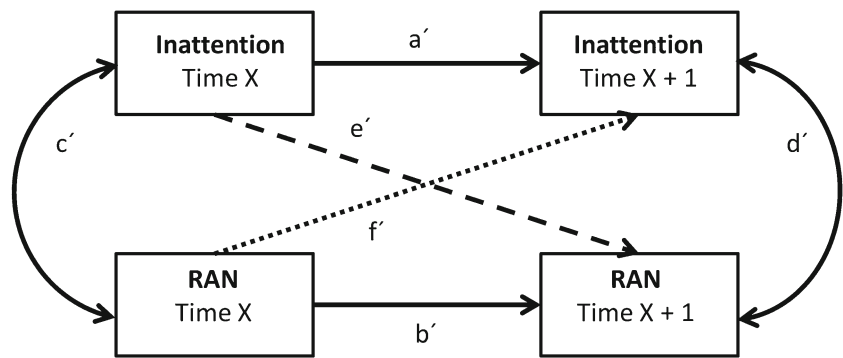
Model #1: Reciprocal model with all paths included
Model #2: RAN predicts inattention; dashed path e' is dropped
Model #3: Inattention predicts RAN; dotted path f' is dropped
Model #4: No cross-lagged associations; paths e' and f' are both dropped

After identifying the best-fitting model in the full sample, secondary analyses were conducted to test if the model fit equally well across gender, country, reading disorder risk, ADHD diagnostic status, and time interval. We hypothesized that the prevailing model would not differ according to gender differences or country, thereby suggesting that the pattern of development would not be due to a gender- or nationality-specific influence, such as social learning. We also hypothesized that RAN would be associated with ADHD inattention behaviors over and above reading skill, so the model would fit equally well for subjects in the bottom 10th percentile for reading skill. Likewise, we expected that the model would fit equally well for children with scores at the extreme ends of the ADHD symptom dimensions. This hypothesis was based on previous research supporting ADHD as an extreme phenotype, rather than a categorical diagnosis (Pennington 2002). In other words, we postulated that the same cognitive processes that influence individual differences in inattention symptoms in unaffected children also explain differences in the extreme symptomatic end of the ADHD symptom dimension.

Furthermore, we tested the equality of cross-lagged path coefficients over time. We predicted that if the reciprocal model fit the data best, we would also see that the magnitude of cross-lag weights in both directions increased over time, supporting a transactional association. In a transactional association, a positive feedback loop would support the reciprocal development of inattention symptoms and RAN, such that over time, each construct would account for more and more variance in the opposite construct. Practically, that would mean that as inattention symptoms worsened, they would decrease RAN speed, which would result in even worse inattention symptoms at the subsequent time point, and so on. In Model 2, RAN has a significant predictive effect on a child's inattention symptoms at each time interval, consistent with the endophenotype theory. In Model 3,

we hypothesized that the effect of inattention symptoms on RAN would be largely experiential; that is, higher inattention symptoms would select children away from experiences that would otherwise reinforce their RAN skills. As environmental demands for RAN increased with more demanding academic settings, children with high inattention symptoms would show increasingly worse RAN and thus the cross-lags from inattention symptoms to RAN would become stronger. The fourth model would support no cross-lagged relations, and would suggest that an unmeasured third variable accounted for the contemporaneous associations between inattention symptoms and RAN.

Finally, the discriminant validity of the developmental pattern between RAN and inattention symptoms was tested by substituting hyperactivity/impulsivity symptoms for inattention symptoms and testing each of the models. A different pattern of results between the RAN and hyperactivity/impulsivity factors would provide further evidence of a unique longitudinal relational pattern between RAN and inattention symptoms.

Results

Preliminary Analyses

Compared to distributions in the overall sample, there were significantly more boys than girls in the 10th percentile for any ADHD subtype ($t=4.89$, $p<0.001$); and significantly fewer Australian subjects fell into the 10th percentile for any ADHD subtype (11 %) than U.S. subjects (18 %; $t=3.446$, $p<0.01$). Subjects in the 10th percentile for any ADHD dimension scored 9.12 standardized points lower on TOWRE reading tests at 2nd grade, on average, than subjects who were in the normal range for ADHD symptoms ($t=8.86$, $p<0.001$).

Birth weight and mean parental education were not significantly different for subjects in the top 10th percentile for any ADHD subtype (birth weight $t=1.379$, $p=0.168$; parent education $t=0.978$, $p=0.328$). ADHD subscale severity scores were in the clinically normal range, overall, in the entire sample (Inattention $M=5.32$, Hyperactivity/Impulsivity $M=4.32$).

Model Comparisons

Inattention and RAN The fits of the four competing models were first compared with Pre-K, K, and 1st grade time points included. Model #1 (reciprocal associations) fit the data best: χ^2 $(4)=22.137$, $p=0.0002$, CFI=0.99, RMSEA= 0.055 (see Table 2). These analyses were then repeated with the U.S. 4th grade data included (in addition to the data from both countries for the first three time points) as a conservative look at whether the previous pattern held through later childhood. Model #1 again fit the data best, with the other models showing a significantly worse fit (see Table 2). However, due to the lack of Australian data at the 4th time point, the remaining path modeling analyses were done only using pre-K, K and 1st grade time points.

We also tested the potential effects of parent education and birthweight on the models, as these demographic values were significantly different for subjects who had missing data at any time point. The results were similar, with Model #1 demonstrating a significantly better fit than the alternative models: χ^2 $(5)=93.34$, $p<0.001$, CFI=0.95, RMSEA=0.108.

Inattention and Residual RAN Given the previously discussed concern about the association between phoneme awareness and RAN, we repeated the model comparisons using a residual RAN factor score that controlled for phoneme awareness, by regressing a phoneme awareness factor score out of RAN at each time point. The results again converged on Model #1 as the best fit χ^2 $(4)=34.254$, $p<0.001$, CFI= 0.980, RMSEA=0.071 (see Table 2).

Group Comparisons Next, we tested the equality of fit for Model #1 across subgroups of our sample to explore the generalizability of these findings. This was done by constraining the cross-lagged and autoregressive path weights to be equal across the subgroups of interest (Muthén and Muthén 1998–2010). Contemporaneous path coefficients were not constrained, as we were only interested in testing whether predictive effects (rather than concurrent ones) were comparable across groups. A non-significant difference in chi-square would indicate that the model fit equally well for both groups and the regression paths did not vary in magnitude across the groups. For all group comparisons, pre-K, K and 1st grade time points were used, with RAN factor scores that did not control for phoneme awareness. Model #1 fit equally well for males versus females ($\Delta\chi^2$ $(12)=4.399$, $p=0.820$); subjects in the 10th percentile for

Table 2 Fit statistics for competing cross-lagged models

Model	χ^2 (df) [a]	CFI	RMSEA	$\Delta\chi^2$ (df) [b]	p
Time Points: Pre-K, K, 1st					
#1 Reciprocal Transactions	22.137 (4)	0.990	0.055		
#2 RAN → INATT	38.904 (6)	0.982	0.060	14.931 (2)	<0.001
#3 INATT → RAN	65.542 (6)	0.967	0.081	38.754 (2)	<0.001
#4 No Cross-lags	85.763 (8)	0.957	0.080	56.860 (4)	<0.001
Time Points: Pre-K, K, 1st and 4th					
#1 Reciprocal Transactions	111.145 (12)	0.961	0.074		
#2 RAN → INATT	127.401 (15)	0.955	0.071	13.325 (3)	<0.01
#3 INATT → RAN	165.942 (15)	0.940	0.082	49.590 (3)	<0.001
#4 No Cross-lags	184.983 (18)	0.934	0.078	63.874 (6)	<0.001
Time Points: Pre-K, K, 1st with RAN Controlled for Phoneme Awareness					
#1 Reciprocal Transactions	34.254 (4)	0.980	0.071		
#2 RAN → INATT	51.663 (6)	0.969	0.071	15.138 (2)	<0.001
#3 INATT → RAN	61.513 (6)	0.963	0.078	24.803 (2)	<0.001
#4 No Cross-lags	79.107 (8)	0.952	0.077	40.012 (4)	<0.001
Time Points: Pre-K, K, 1st with Hyperactivity/Impulsivity as Behavioral Factor					
#1 Reciprocal Transactions	31.524 (4)	0.984	0.068		
#2 RAN → ATT	33.728 (6)	0.984	0.055	1.869 (2)	0.393
#3 ATT → RAN	32.645 (6)	0.984	0.054	0.941 (2)	0.625
#4 No Cross-lags	34.925 (8)	0.984	0.047	2.870 (4)	0.580

[a] All models were significant at $p<0.001$.

[b] Chi-square difference calculated using the Satorra-Bentler Scaled Chi-Square.

4th grade data only included U.S. subjects.

ADHD-inattention or combined symptoms versus subjects in the normal range ($\Delta\chi^2$ (12)=10.753, p=0.216); and a group at risk for RD versus a group in the normal range ($\Delta\chi^2$ (12)=9.127, p=0.332; Table 3).

Country comparisons demonstrated a significantly different fit for U.S. versus Australian samples (see Table 3). Further analyses, in which we systematically constrained one group of path weights at a time, revealed that this fit difference was driven entirely by a difference in the autoregressive path from pre-K inattention to Kindergarten inattention scores. Stability of inattention at that initial time interval was significantly higher for the Australian sample (B[SE]=1.046[0.141]) than the U.S. sample (B[SE]=0.116[0.027]).

Path Weights

We examined the values of the autoregressive, contemporaneous, and cross-lagged path coefficients for further interpretation of the size of the relationships across these varying paths, using the pre-K to 1st grade model. Path coefficients with $p<0.01$ were considered statistically significant. Autoregressive paths for the inattention symptoms factor revealed adequate longitudinal stability of this factor when variance attributable to contemporaneous and cross-lagged associations were simultaneously being accounted for: unstandardized B (SE)=0.161(0.03) and 0.527(0.03). However, longitudinal stability of inattention symptoms from K-1st grade was significantly higher than that from Pre-K to K, as evidenced by a significant chi square change when the autoregressive paths were constrained to be equal over time: $\Delta\chi^2$ (1)=74.22, $p<0.001$. Autoregressive, unstandardized path weights for RAN were 0.536 (0.01) and 0.788 (0.03), and were likewise significantly different across time intervals: $\Delta\chi^2$ (1)=45.705, $p<0.001$.

The contemporaneous path weights between inattention symptoms and RAN decreased steadily over time, with unstandardized values of B (SE)=0.023(0.00), $p<0.001$ at Pre-K and 0.006(0.00), $p<0.01$ at 1st grade. The model had a significantly worse fit when the contemporaneous path weights were constrained to be equal:

$\Delta\chi^2$ (2)=35.956, $p<0.001$, indicating that these path coefficients were significantly different overall.

Unstandardized cross-lag path weights from RAN to inattention symptoms were significant at $p<0.01$, with path coefficients B (SE)=1.821 (0.35) at pre-K to K, and 1.309 (0.30) at K to 1st grade; they were not significantly different from one another ($\Delta\chi^2$ (1)=0.677, $p=0.411$), indicating that the direct influence of RAN on inattention was stable over time. The cross-lag effects from inattention to RAN were likewise comparable across time intervals, but only the path from pre-K to K was significant at $p<0.01$. Unstandardized values were B (SE)=0.007 (0.00) at pre-K to K, and 0.005 (0.00) between K and 1st grade. We next evaluated relative predictive power of each construct on the other. The standardized cross-lag path coefficients were constrained to be equal in each direction, one time point at a time, to test for differences in the cross-lag effects. It was necessary to constrain standardized coefficients (rather than unstandardized) to control for the difference in variance between the two constructs. The prospective effect of RAN on inattention symptoms was not significantly higher than paths in the opposite direction for predictions of Kindergarten and 1st grade outcomes (Table 4).

Path weights did not change when the 4th grade time point was included in the model, so standardized path coefficients are recorded for all four time points in Fig. 2. As described above, paths were not significantly different when we constrained across gender, ADHD-I/C and No ADHD, or RD risk and no RD risk. Standardized cross-lagged coefficients for each subgroup are recorded in Table 5.

Discriminant Validity Analysis

To determine whether the longitudinal association between RAN and inattention was specific to inattention rather than representative of a general association between RAN and ADHD, we tested the model fit with hyperactivity/impulsivity. We tested all four models using three time points, and a residual hyperactivity/impulsivity factor score that controlled for inattention. With hyperactivity/impulsivity as the

Table 3 Model #1 fit statistics across demographics

Model	χ^2 (df)	CFI	RMSEA	$\Delta\chi^2$ (df) [a]	p
Country	69.238 (16), $p<0.001$	0.972	0.660	30.910 (8)	<0.01
Gender	33.076 (16), $p<0.01$	0.991	0.038	4.399 (8)	0.820
ADHD I/C	37.160 (16), $p<0.01$	0.986	0.043	10.753 (8)	0.216
RD Risk	36.412 (16), $p<0.01$	0.987	0.043	9.127 (8)	0.332

[a] Chi-square difference calculated using the Satorra-Bentler Scaled Chi-Square.

Time points included pre-K, K and 1st grade. ADHDI/C subjects were in the 10th percentile for ADHD-I or ADHD-C; RD Risk subjects were in the lowest 10th percentile for TOWRE subtest standard scores

Table 4 Model #1 fit statistics with standardized cross-lag path constraints

Time Interval	χ^2 (df) [a]	CFI	RMSEA	$\Delta\chi^2$ (df) [b]	p
Pre-K – K	25.070 (5)	0.989	0.052	2.448 (1)	0.118
K – 1st	25.815 (5)	0.988	0.053	3.859 (1)	0.050

[a] All models were significant at $p<.001$

[b] Chi-square difference calculated using the Satorra-Bentler Scaled Chi-Square

behavioral construct, all four models fit equally well. None of the cross-lag path weights were statistically significant (Fig. 3). In concordance with the principle of parsimony, Model #4, which assumes no cross-lagged relationships between hyperactivity/impulsivity and RAN, has the best fit χ^2 (8) =34.93, $p<0.001$, CFI=0.984, RMSEA=0.047 (Table 2). Thus, as we proposed, the unidirectional, longitudinal association between RAN and inattention symptoms was specific to these two constructs.

Discussion

The goal of this study was to clarify the associations between ADHD inattention symptoms and RAN over the course of early childhood. The best fitting model of these associations suggests that individual differences in rapid naming speed and inattention symptoms in early childhood have a reciprocal predictive relationship, whereby each construct explains variance in the other at subsequent time points. This model is consistent with a reciprocal theory of development between RAN and inattention symptoms. In this theory, inattentive behavior leads to worse rapid naming speed (e.g. by the child selecting away from activities that would enhance rapid naming speed), while poor rapid naming speed (which is highly correlated with processing speed)

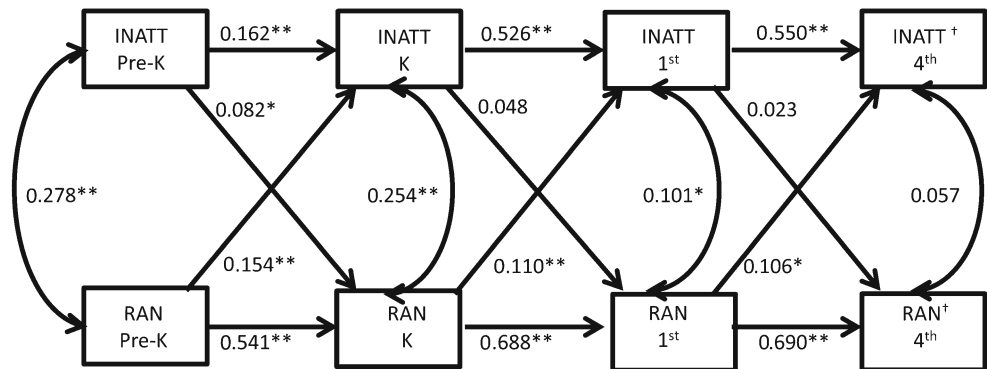**Table 5** Standardized cross-lagged coefficients by group and time interval

Subgroup	RAN→Inattention		Inattention→RAN	
	Pre-K to K	K to 1st	Pre-K to K	K to 1st
Boys	0.155**	0.113*	0.077	0.052
Girls	0.156**	0.111*	0.091*	0.044
ADHDI/C	0.112*	0.120**	0.059	0.019
No ADHD	0.100*	0.098**	0.055	0.026
RD Risk	0.115	0.143	−0.102	0.114
No RD Risk	0.144**	0.065*	0.063	0.039

**$p<0.001$, *$p<0.01$

is predictive of worse inattention. The reciprocal model fit the data well regardless of gender, nationality, or RD risk, and fit equally well among children at the 10th percentile for ADHD inattention or combined symptoms.

Notably, the cross-lagged effects in both directions were consistent from Pre-K through 1st grade, suggesting that RAN has an ongoing, stable effect on inattention symptoms, and vice versa. This pattern of stability is not entirely consistent with a transactional model, because in a transactional model we would expect the positive feedback loop to lead to an increase in effect size over time. Further, although the cross-lagged effects for RAN symptoms predicting inattention look larger in magnitude than those in the opposite direction, we tested the equality of the standardized cross-lag coefficients at each time interval, and they were not significantly different. Therefore, all results converge on the reciprocal, saturated model as the best fit. Importantly, these results do not appear to depend on either the comorbidity of ADHD with reading disability, associated hyperactivity/impulsivity symptoms, or the phonological component of rapid naming measures, as the results were unchanged when these factors were controlled. Instead,

Fig. 2 Cross-lagged path model of inattention and rapid naming speed factors

**$p<.001$, *$p<.01$, [†]Data from the U.S. sample only. Path coefficients are standardized.
INATT=ADHD Inattention symptom factor score; RAN=Rapid naming factor score.

Fig. 3 Cross-lagged path model of hyperactivity/impulsivity and rapid naming speed factors
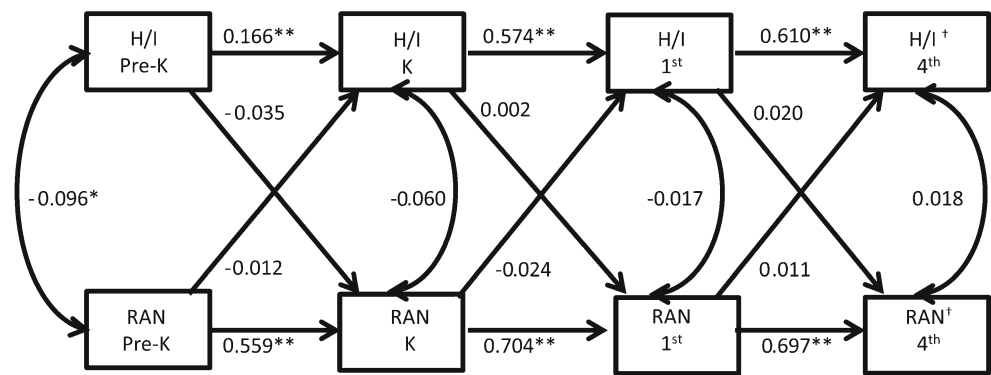
**p<.001, *p<.01, †Data from the U.S. sample only. Path coefficients are standardized.
H/I=ADHD hyperactivity/impulsivity symptom factor score; RAN=Rapid naming factor score.
Note: the best fitting model with H/I and RAN was Model #4, which included no cross-lagged paths.

these results support a fairly specific, reciprocal relation between the processing speed aspect of rapid naming and the inattention symptom dimension of ADHD.

The observed pattern of results has several important implications for theoretical models of ADHD. The fact that RAN predicts the development of inattention symptoms is consistent with previous theories that proposed that naming or processing speed may be a useful cognitive endophenotype for genetic studies of ADHD. Potential explanations for this observed association include, 1) slow RAN is an indicator of slow cognitive speed, and cognitive speed directly supports attentive behaviors; or 2) RAN is an index of a neurobiological trait, such as white matter integrity, that is necessary for attentive behaviors. In addition, these results suggest that RAN may be useful for early detection of children at risk for ADHD-Inattentive or -Combined types in a clinical setting. It follows that RAN or nonverbal processing speed tasks may also be useful outcome measures for intervention studies. In contrast, the small but significant associations between inattention and later RAN in early childhood suggest that inattention may influence experiential learning at an early age, resulting in differential effects on rapid naming skills, at least from pre-K to Kindergarten. This has implications for potential cognitive benefits of behavioral interventions for ADHD.

Finally, the fact that the model fit equally well for children in the 10th percentile for ADHD inattentive and combined symptom severity is consistent with the hypothesis that similar neuropsychological processes underlie the development of inattention symptoms at the extreme end of functioning, as well as within the normal range.

Discriminant validity for our model was supported by a comparison of the model fit when hyperactivity/impulsivity was included as the behavioral factor in place of inattention. With hyperactivity/impulsivity as the behavioral factor, the best model fit included no cross-lagged associations. This

finding indicated that the reciprocal relationship between RAN and inattention factor scores was unique, and it is consistent with previous literature that has found fewer cognitive correlates of the DSM-IV ADHD Predominantly Hyperactive/Impulsive subtype diagnosis. Our results suggest that the cognitive endophenotypes that predict inattention in children do not necessarily predict hyperactivity/impulsivity, and therefore may reflect different underlying cognitive processes that influence these distinct behavioral symptoms. Further, hyperactive/impulsive behaviors may not result in the same experiential processes that select children away from activities that support development of RAN.

Limitations and Future Directions

Despite the consistency of our findings, the results of this study should be interpreted with some caveats. First, although Model #1 had the best fit, the cross-lag effects from inattention to RAN were small, and only statistically significant at the pre-K to Kindergarten time interval. Thus, although the current results indicate that RAN and inattention symptoms demonstrate reciprocal, longitudinal associations in early elementary school, additional research is needed to clarify whether this pattern remains consistent in different samples and at later time points.

It is important to emphasize that the associations we found are not necessarily causal. The results are consistent with a causality hypothesis, in that RAN precedes and predicts inattention above and beyond autoregressive effects. However, the variation in measures of inattention and RAN across time points limits our ability to make a causal conclusion. Further, in order to determine causality we would need to be able to test all possible confounders that could be driving the association between RAN and inattention, and we would need to manipulate performance

on both constructs. A confounder would necessarily be highly associated with RAN scores and could, in theory, be the true cause of variation in inattention symptoms. For example, we hypothesized that phoneme awareness, which is correlated with RAN, could possibly contribute to the association between RAN and attention. Additional potential confounders that we did not measure might be general intelligence, exposure to reading, and school environment. Future work should attempt to replicate the model using a measure of motoric processing speed as the cognitive factor, as well as alternate cognitive factors that have been proposed as endophenotypes for ADHD, such as inhibition, response time, and working memory.

A minor limitation of this study was the use of a twin sample. Although twins are generally 3–4 weeks premature and develop language more slowly in early childhood, twins have not been found to show important differences from singletons in the domain of psychopathology (Plomin, DeFries, McClearn, and Rutter 1997). Additionally, the use of twins will allow us to conduct future behavioral genetic analyses to test if common genes influence the reciprocal, longitudinal associations between RAN and inattention symptoms. Nonetheless, future research should replicate the analyses presented here using singletons, in order to test whether the current findings replicate in a non-twin sample.

As this was a population sample and clinical diagnostic information was not collected, we were unable to account for the use of stimulant (or other) medication in our subjects. Use of medication to control ADHD behaviors might affect the parent- and tester-ratings of children who would otherwise display stronger inattention symptoms. However, the stability of the inattention construct suggests that this was not a major confounding factor in this sample. Rates of pharmaceutical treatment of ADHD increases with age, so data at the later time points were likely to be most affected by this unknown confounder.

At the pre-K time point, only DBRS ratings by parent and tester were used to measure symptoms of inattention and hyperactivity/impulsivity, while at all other time points, parent ratings on the SWAN were also included. SWAN measurements provide additional variance in this construct and thus may account for the fact that the autoregressive path for inattention from pre-K to K was weaker than subsequent paths. Likewise, RAN subtests were not the same at all time points. Although the strength of the longitudinal stability of the RAN factor suggests they were measuring the same construct, this difference may be considered a minor limitation in the study. Future work would benefit from using identical measures of both inattention and RAN at each time point.

Finally, there exists a possibility that significant transactions between inattention and RAN took place prior to

entrance into the school environment. Future work would benefit from inclusion of very young children in this model.

Conclusions

The current study supports a reciprocal model of the development of inattention symptoms and rapid naming skills in early childhood. In this model, poor RAN at an early age predicts development of subsequent inattention symptoms, and we hypothesize that this is due to cognitive processes that underlie attention and focus behaviors. Likewise, the model supports small effects in the opposite direction, whereby inattentive behavior at a young age predicts worse rapid naming skills in later years. The effect of inattention on RAN is hypothesized to be due to self-selection of experiences and environment; in other words, children with low attention are less likely to engage in activities that would foster development of RAN, and vice versa. According to the model, self-selection of experiences is itself influenced by previous development of rapid naming skills. Importantly, the findings in this study also offer further support of an extreme phenotype model of ADHD, whereby inattention symptoms exist on a continuum, and the underlying cognitive processes that affect inattention remain constant regardless of where an individual falls on that continuum. Implications of this research for clinical practice include the use of rapid naming tasks in childhood evaluations for ADHD to predict behavioral outcomes at an early age, as well as offer external validity to behavioral observations. Interventions for ADHD might focus on improving cognitive speed, and research on behavioral interventions for ADHD may consider measuring progress at the neuropsychological level with rapid naming speed scores. Previous research has documented improvements in color naming speed in response to stimulant medications (Tannock et al. 2000); thus, future longitudinal work could benefit from the inclusion of stimulant medication as a covariate in the models. Additionally, replication of the current findings should be attempted using additional ages and alternate behavioral, genetic and cognitive factors.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (Fourth Edition-TR ed.). Washington, D.C.: American Psychiatric Association.

Arnett, A. B., Pennington, B. F., Friend, A., Willcutt, E. G., Byrne, B., Samuelsson, S., & Olson, R. K. (2011). The SWAN captures variance at both the negative and positive ends of the ADHD symptom dimension. *Journal of Attention Disorders*. doi:10.1177/1087054711427399.

Barkley, R. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*(1), 65.

Barkley, R., & Murphy, K. (1998). *Attention deficit hyperactivity disorder: A clinical workbook* (Secondth ed.). New York: Guilford Press.

Berlin, L., Bohlin, G., & Rydell, A. (2003). Relations between inhibition, executive functioning, and ADHD symptoms: A longitudinal study from age 5 to 8 1/2 years. *Child Neuropsychology, 9*(4), 255.

Bidwell, L. C., Willcutt, E., DeFries, J. C., & Pennington, B. F. (2007a). Testing for neuropsychological endophenotypes in siblings discordant for attention-deficit/hyperactivity disorder. *Biological Psychiatry, 62*(9), 991.

Bidwell, L. C., Willcutt, E. G., DeFries, J. C., & Pennington, B. F. (2007b). Testing for neuropsychological endophenotypes in siblings discordant for attention-deficit/hyperactivity disorder. *Biological Psychiatry, 62*, 991–998.

Brocki, K. (2006). Developmental change in the relation between executive functions and symptoms of ADHD and co-occurring behaviour problems. *Infant and Child Development, 15*(1), 19.

Brocki, K., Nyberg, L., Thorell, L. B., & Bohlin, G. (2007). Early concurrent and longitudinal symptoms of ADHD and ODD: Relations to different types of inhibitory control and working memory. *Journal of Child Psychology and Psychiatry, 48*(10), 1033.

Byrne, B. B., Delaland, C., Fielding-Barnsley, R., Quain, P., Samuelsson, S., Hoien, T., & Olson, R. K. (2002). Longitudinal twin study of early reading development in three countries: Preliminary results. *Annals of Dyslexia, 52*(1), 47–73.

Carte, E. T., Nigg, J. T., & Hinshaw, S. P. (1996). Neuropsychological functioning, motor speed, and language processing in boys with and without ADHD. *Journal of Abnormal Child Psychology, 24*(4), 481–498.

Castellanos, F. X., & Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: The search for endophenotypes. *Nature Reviews Neuroscience, 3*(8), 617.

Chhabildas, N. N., Pennington, B. F., & Willcutt, E. G. (2001). A comparison of the neuropsychological profiles of the DSM-IV subtypes of ADHD. *Journal of Abnormal Child Psychology, 29*(6), 529–540.

Hartman, C., Rhee, S., Willcutt, E. G., & Pennington, B. F. (2007). Modeling rater disagreement for ADHD: Are parents or teachers biased? *Journal of Abnormal Child Psychology, 35*(4), 536–542.

Hays, R. D., Marshall, G. N., Wang, E. Y. I., & Sherbourne, C. D. (1994). Four-year cross-lagged associations between physical and mental health in the medical outcomes study. *Journal of Consulting and Clinical Psychology, 62*(3), 441–449.

Hinshaw, S. P., Owens, E. B., Sami, N., & Fargeon, S. (2006). Prospective follow-up of girls with attention-Deficit/Hyperactivity disorder into adolescence: Evidence for continuing cross-domain impairment. *Journal of Consulting and Clinical Psychology, 74*(3), 489–499.

Kalff, A. C., De Sonneville, L. M. J., Hurks, P. P. M., Hendriksen, J. G. M., Kroes, M., Feron, F. J. M., & Jolles, J. (2005). Speed, speed variability, and accuracy of information processing in 5 to 6-year-old children at risk of ADHD. *Journal of the International Neuropsychological Society, 11*, 173–183.

Kline, R. B. (2005). *Principles and practice of structural equation modeling (Second* (Editionth ed.). New York: Guilford Press.

Lahey, B. B., & Willcutt, E. (2010). Predictive validity of a continuous alternative to nominal subtypes of attention-deficit hyperactivity disorder in DSM-IV. *Journal of Clinical Child and Adolescent Psychology, 39*(6), 761–775.

Lahey, B. B., Pelham, W. E., Stein, M. A., Loney, J., Trapani, C., Nugent, K., & Baumann, B. (1998). Validity of DSM-IV attention-deficit/hyperactivity disorder for younger children. *Journal of the American Academy of Child and Adolescent Psychiatry, 37*(7), 695.

Lahey, B. B., Pelham, W. E., Loney, J., Kipp, H., Erhardt, A., Lee, S. S., & Massetti, G. (2004). Three-year predictive validity of DSM-IV attention deficit hyperactivity disorder in children diagnosed at 4–6 years of age. *The American Journal of Psychiatry, 161*, 2014–2020.

Lahey, B. B., Pelham, W. E., Loney, J., Lee, S. S., & Willcutt, E. (2005). Instability of the DSM-IV subtypes of ADHD from preschool through elementary school. *Archives of General Psychiatry, 62*(8), 896–902.

Larsson, H., Lichtenstein, P., & Larsson, J. O. (2006). Genetic contributions to the development of ADHD subtypes from childhood to adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry, 45*(8), 973.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130–149.

McGrath, L., Shanahan, M., Santerre-Lemmon, L., Barnard, H., Willcutt, E., Olson, R., Pennington, B. (2011). A multiple deficit model of reading disability and attention-deficit/hyperactivity disorder: Searching for a shared cognitive deficit. *Journal of Child Psychology and Psychiatry*, 557.

Muthén, B. K., & Muthén, B. O. (1998–2010). *Mplus user's guide* (Sixth Edition ed.). Los Angeles, CA: Muthén & Muthén.

Nigg, J. T. (2005). Attention, task difficulty, and ADHD. *British Journal of Developmental Psychology, 23*(4), 513.

Nigg, J. T., Blaskey, L. G., Huang-Pollock, C. L., & Rappley, M. D. (2002). Neuropsychological executive functions and *DSM-IV* ADHD subtypes. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(1). doi:10.1097/00004583-200201000-00012.

Pennington, B. F. (2002). *The development of psychopathology: Nature and nurture*. New York: Guilford Press.

Plomin, R., DeFries, J. C., McClearn, G. E., & Rutter, M. (1997). *Behavioral genetics*. New York: Freeman.

Samuelsson, S., Byrne, B., Quain, P., Wadsworth, S., Corley, R., DeFries, J. C., & Olson, R. (2005). Environmental and genetic influences on prereading skills in australia, scandinavia, and the united states. *Journal of Educational Psychology, 97*(4), 705–722.

Satcher, D. (2000). Mental health: A report of the surgeon General—executive summary. *Professional Psychology: Research and Practice, 31*, 5–13.

Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H. Heijmans, D. S. G. Pollock, & A. Satorra (Eds.), *Innovations in multivariate statistical analysis. A festschrift for heinz neudecker* (pp. 233–247). London: Kluwer Academic Publishers.

Shanahan, M., Pennington, B. F., Yerys, B. E., Scott, A., Boada, R., Willcutt, E. G., & DeFries, J. C. (2006). Processing speed deficits in attention Deficit/Hyperactivity disorder and reading disability. *Journal of Abnormal Child Psychology, 34*, 585–602.

Sonuga-Barke, E. J. (2005). Causal models of attention-Deficit/Hyperactivity disorder: From common simple deficits to multiple developmental pathway. *Biological Psychiatry, 57*(11), 1231.

Swanson, J., Shuck, S., Mann, M., Carlson, C., Hartman, K., Sergeant, J., & McCleary, R. (2006). *Categorical and dimensional definitions and evaluations of symptoms of ADHD: The SNAP and SWAN rating scales. Irvine*. Irvine: University of California.

Tannock, R. R., Martinussen, R., & Frijters, J. (2000). Naming speed performance and stimulant effects indicate effortful, semantic processing deficits in attention-deficit/hyperactivity disorder. *Journal of Abnormal Child Psychology, 28*(3), 237–252.

Torgesen, J., Wagner, R., & Rashotte, C. (1999). *A test of word reading efficiency (TOWRE)*. Austin: PRO-ED.

Vukovic, R. K., & Siegel, L. S. (2006). The double-deficit hypothesis: A comprehensive analysis of the evidence. *Journal of Learning Disabilities, 39*, 25–47.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *The comprehensive test of phonological processes (CTOPP)*. Austin: PRO-ED.

Wåhlstedt, C., Thorell, L. B., & Bohlin, G. (2008). ADHD symptoms and executive function impairment: Early predictors of later behavioral problems. *Developmental Neuropsychology, 33*(2), 160.

Weiler, M. D., Bernstein, J. H., Bellinger, D. C., & Waber, D. P. (2000). Processing speed in children with attention Deficit/Hyperactivity disorder, inattentive type. *Child Neuropsychology, 6*(3), 218–234.

Willcutt, E. G., Pennington, B. F., Olson, R. K., Chhabildas, N., & Hulslander, J. (2005). Neuropsychological analyses of comorbidity between reading disability and attention deficit hyperactivity disorder: In search of the common deficit. *Developmental Neuropsychology, 27*(1), 35–78.

Willcutt, E. G., Betjemann, R. S., Wadsworth, S. J., Samuelsson, S., Corley, R., DeFries, J. C., & Olson, R. K. (2007). Preschool twin study of the relation between attention-deficit/hyperactivity disorder and prereading skills. *Reading and Writing, 20*, 103–125.

Willcutt, E. G., Nigg, J. T., Pennington, B. F., Solanto, M. V., Rohde, L. A., Loo, S., Lahey, B. B. (2012). *Validity of DSM-IV attention-deficit/hyperactivity disorder symptom dimensions and subtypes.* Unpublished manuscript.