Published in final edited form as:

J Atten Disord. 2011 October ; 15(7): . doi:10.1177/1087054709356379.

Response Time Variability is Related to Parent Ratings of Inattention, Hyperactivity and Executive Function

Lorena Gómez-Guerrero, M.D., Cristina Domínguez-Martín, M.D., María Angeles Mairena, Cl.Psy., Adriana Di Martino, M.D., Jing Wang, M.S, Alan Mendelsohn, M.D., Benard Dreyer, M.D., Peter Isquith, Ph.D., Gerard Gioia, Ph.D., Eva Petkova, Ph.D., and F. Xavier Castellanos, M.D.

Abstract

Objective: Individuals with Attention-Deficit/Hyperactivity Disorder (ADHD) are often characterized as inconsistent across many contexts. ADHD is also associated with deficits in executive function. We examined the relationships between response time variability on five brief computer tasks to parents' ratings of ADHD-related features and executive function in a group of children with a broad range of ADHD symptoms from none to full diagnosis.

Methods: We tested 98 children (mean age 9.9 ± 1.4 yrs; 66 boys) from community clinics on short Tasks of Executive Control (TEC) and the Eriksen Flanker Task, while a parent completed the Conners Parent Rating Scale and Behavior Rating Inventory of Executive Function.

Results: Variability for two of the TEC tasks explained significant proportions of the variance of all five ADHD-related Conners subscales and several executive function subscales. By contrast, variability on the flanker task or mean response times for any task were not associated with any rating scale.

Conclusions: The significant dimensional relationships observed between variability measures and parent ratings support the utility of response time variability as an objective measure in ADHD and aspects of executive functioning that is superior to response time means or accuracy measures.

Keywords

Attention-Deficit/Hyperactivity Disorder; Intra-subject Variability; Rating Scales; Executive Function

Attention-Deficit/Hyperactivity Disorder (ADHD) is the most studied and prevalent behavioral disorder affecting school-age children (Lahey et al., 1994). The diagnosis is based exclusively on the psychiatric history, as objective measures are not sufficiently sensitive and specific (Castellanos, 1997). Dimensional rating scales, which are inexpensive and easily administered, are used to support diagnostic assessment, index severity, and track benefits from treatment. One of the most commonly used dimensional rating scales for ADHD is the Conners Parent Rating Scale-Revised (CPRS; Conners, 1997). However, while rating scales quantify symptoms, they cannot inform models of the underlying pathophysiology. By contrast, classical neuropsychological measures such as the Continuous Performance task were designed to link clinical behavioral deficits to the effects of gross brain lesions (Rosvold et al., 1956). A wide range of neuropsychological tests have been

studied extensively in relation to ADHD, with a particular focus on measuring executive functions (Nigg, 2005). Statistically significant differences between groups of children with ADHD and healthy comparisons with respect to executive function measures have been consistently reported (Willcutt et al., 2005). However, the effect sizes are typically small to moderate, indicating that such measures cannot contribute diagnostically (Happé et al., 2006; Sonuga-Barke et al., 2008). Despite the widespread use of neuropsychological tasks for assessing executive function and examining between-group differences, the cognitive deficits associated with ADHD remain underspecified (Castellanos et al., 2006).

Castellanos and Tannock (2002) drew attention to the ubiquity of behavioral and neuropsychological inconsistency in ADHD and suggested that it represents a hallmark of the disorder. Behavioral inconsistency can be operationalized in terms of intra-subject variability (ISV), which is easily measured with response time (RT) tasks (Castellanos et al., 2005). The conclusion that variability is a better summary of individual performance in ADHD than measures of accuracy or mean RT was supported by Klein et al. (2006). Those authors compared 57 children with ADHD to 53 age- and sex-matched controls on a Continuous Performance task, a Go/No-Go task, a Stop-Signal task, and an N-Back task. They found that ISV (indexed as RT standard deviation (RT-SD)) comprised a single principal component in ADHD across the four tasks. In other words, children with ADHD who were highly variable on one task tended to be highly variable on the remaining tasks. Consequently, group differences in ISV produced substantially larger effect sizes than mean RT, measures of inhibitory control, or of accuracy (Klein et al., 2006). Additionally, ISV has been found to be heritable in ADHD and in controls, in contrast to measures of executive function or delay aversion (Kuntsi & Stevenson, 2001; Kuntsi et al., 2006; Andreou et al., 2007). The significant heritability of ISV in ADHD indicates that it is influenced by specific genetic factors, which are hypothesized to underlie both objective behavioral inconsistency and the clinical syndrome of ADHD itself.

The increasing recent focus on variability in ADHD poses questions regarding the suitability of specific cognitive tasks, specifically with respect to task duration. In general, longer tasks tend to be experienced as more boring than shorter ones and are more difficult to administer to large samples. However, tasks that are too brief may be insufficiently reliable to yield valid results. Studies reporting a significant relationship between ISV and ADHD have used tasks ranging from 5.5 min (Johnson et al., 2007) to 15 min (Di Martino et al., 2008). However, a study using a simple choice response task lasting 182 s (~ 3 min) failed to find a significant group difference in ISV between 53 children with ADHD and 85 typically developing controls (Geurts et al., 2008). Motivated to find neuropsychological tests that could relate ISV to ADHD and that could be administered in future large-scale studies, we selected a battery of short easily administered computer tasks (< 3.5 min each). The performance battery under investigation consisted of the Tasks of Executive Control (TEC) – a recently developed battery designed to assess two key components of Executive Function, working memory and inhibitory control, in children between ages 5 and 18 (Isquith et al., 2009). Additionally, we included a short version of an Eriksen Flanker Task (EFT), which we previously used to examine RT-ISV in children with ADHD in longer versions (Scheres et al., 2001; Castellanos et al., 2005; Di Martino et al., 2008).

The principal purpose of this pilot study was to determine whether RT-ISV during short, easily administered computer tasks would be related to CPRS parent ratings of ADHD. Instead of comparing ISV in subjects with and without a diagnosis of ADHD as we had previously done (Castellanos et al., 2005; Di Martino et al., 2008), we used dimensional measures related to ADHD to estimate percent explained variation in those scales by ISV. This approach is based on a perspective of ADHD as the end of a continuum of functioning rather than a discrete disorder (Levy et al., 1997). A dimensional approach facilitates large

scale investigations and enhances statistical power because it maximizes the number of participants who can be included. Our primary measure of ISV is the coefficient of variation of response time – ($RTCV = SD/mean\ RT$), a unit-less measure that can be compared across tasks and populations (Klein et al., 2006). As dimensional measures of ADHD, we selected the three global DSM-IV ADHD CPRS subscales and the two CPRS primary subscales most related to the core symptoms of ADHD (*Cognitive Problems/Inattention* and *Hyperactivity*).

Given the prominence of the executive dysfunction hypothesis of ADHD (Willcutt et al., 2005), and our use of tasks designed to quantify aspects of Executive Function, we also explored the relation between RTCV and parent ratings of executive dysfunction. Specifically, we used the Behavior Rating Inventory of Executive Function-Parent Form (BRIEF; Gioia et al., 2000) as a measure of the impact of executive function deficits on daily life. Recent investigations have found that parent and teacher BRIEF ratings contribute significantly to the diagnosis of ADHD (Toplak et al., 2008; Gioia et al., 2002) although some studies failed to detect a relationship between BRIEF scores and performance based measures of executive function (Anderson et al., 2002). For this exploratory aim, we included all the BRIEF subscales.

The primary study hypothesis was that RTCV would explain significant proportions of the variance in parent ratings related to ADHD in contrast to mean RT, which would not explain significant proportions of these variances. Second, although we expected that standard task performance measures (accuracy, omissions, and directional errors) would also be significantly related to the selected Conners ratings, we predicted that RTCV would still explain a significant proportion of the remaining variability.

METHOD

Participants

The study was approved by the institutional review board of New York University School of Medicine and by the Research Committee of Bellevue Hospital Center. Subjects were recruited from general pediatric clinics at Bellevue Hospital Center and from the Roberto Clemente Family Therapy Center in New York City. Parents of all children between the ages of 8 and 12 years who were visiting these clinics during the recruitment period were given initial information about the study, asked if they were willing to participate, and screened, mainly by phone. Signed informed consent, in English or Spanish as preferred, was obtained from a parent, and informed assent was obtained from each child at the in-person visit, before any research data were obtained in the preferred language. Our bilingual approach allowed this study to be conducted in a sample of mainly Latino children, who tend to be underrepresented in research studies (Bauermeister et al., 2003) and who represent the majority ethnic group in these two institutions. Families received a \$25 bookstore gift certificate per child for participation.

A total of 102 children were enrolled. One child was excluded due to parent report of neurological problems and three children were excluded due to incomplete parent questionnaires. Thus, the study sample consisted of 98 children (66 males) between the ages of 8 and 12 years (mean 9.9, SD 1.4 years) from 73 families. Presence or absence of ADHD diagnosis was not a part of the inclusion/exclusion criteria. Thus, children with a wide spectrum of ADHD symptoms were included, ranging from none or a few to many symptoms, as reported by a parent. Normal cognitive functioning and intelligence sufficient to understand the study tasks were required and assessed informally based on parent report of educational placement. Exclusion criteria included parent report or evidence of sensory and motor dysfunction, autism spectrum disorder, psychosis or intellectual deficiency.

Materials and Procedure

Participation involved a single visit. If a child was being treated with a stimulant, their parent was asked to withhold the stimulant on the day of the scheduled session (n=11). Symptoms related to ADHD were assessed by the *CPRS Long Version* (Conners, 1997) a widely used standardized instrument containing 14 subscales. We selected five CPRS subscales to test our hypothesis. As summary measures, we chose the three DSM-IV CPRS subscales (*DSM-IV: Inattentive*, *DSM-IV: Hyperactive-Impulsive* and *DSM-IV: Total*). Additionally, we were interested in the relationship between RTCV and the two CPRS primary subscales related to two sets of core symptoms of ADHD (*Cognitive Problems/Inattention* and *Hyperactivity*).

Parents also completed the *BRIEF-Parent Form* (Gioia et al., 2000) to assess the everyday behavioral manifestations of children's executive control functions. The BRIEF is composed of 86 items rated on a three-point Likert scale (Never, Sometimes, Often). Subscales reflect eight subdomains of executive function. The *Inhibit*, *Shift*, and *Emotional Control* subdomains are aggregated to yield the composite *Behavioral Regulation Index* (BRI). The subdomains *Initiate*, *Working Memory*, *Plan/Organize*, *Organization of Materials*, and *Monitor* are combined into a composite Metacognition Index (MI). The BRI and MI are in turn combined to yield an overall Global Executive Composite (GEC). We chose an exploratory approach and tested our hypothesis for all the BRIEF subscales. CPRS and BRIEF scores are reported as T-scores.

Parents also completed questionnaires about demographic, educational and basic clinical history regarding their child, including prior diagnoses. While a parent was providing the ratings, their child was tested on a battery of five computer tasks. The computer battery was administered on the same dedicated laptop to all children. The administration of instructions, practice trials and study trials were overseen by LGG or CDM.

Tasks of Executive Control—Task order was the same for all children. The first four were part of the Tasks of Executive Control (TEC), designed to assess working memory and inhibitory control in children between ages 5 and 18 (Isquith et al., 2009). The first task was a choice response task without inhibition (Zero-Back, 0B); the second task added a Go/No Go inhibitory component (Zero-Back with Inhibition, 0BI). This was followed by a One-Back task (1B); the fourth was a One-Back task with inhibition (1BI).

In the basic 0B task, participants are asked to sort all non-zebra objects appearing in the center of the screen into a blue box; zebras are to be sorted into the red box. Sorting is done by pressing the relevant shift key, identified with a blue or red dot, with the corresponding right or left hand. 0BI includes the inhibitory requirement that when an object is visually framed in a box, the participant must refrain from pressing either red or blue buttons. The One-Back tasks, with or without the inhibition rule, are variations on the preceding tasks with an additional constraint. Instead of sorting on the basis of a target such as zebras, participants are asked to place an object in the red box if it appears twice in a row. Participants practiced all tasks and received corrective feedback before proceeding. Stimuli were presented for 400 ms and the inter-stimulus interval varied randomly between 1350, 1600 and 1850 ms. Each task lasted approximately 3 min and 25 s, and was followed by a brief break (< 1 min).

Eriksen Flanker Task—The last task administered consisted of an arrow version of the Eriksen Flanker task (EFT) (Scheres et al., 2001) lasting three min. Target stimuli were arrows pointing to the right or to the left, presented at the center of the screen. The target stimulus was surrounded by four flanking distracters, two on each side. Three trial types were presented: neutral, congruent and incongruent. Neutral trials consist of the central

target arrow flanked by rectangles. In congruent trials, all five arrows point in the same direction. In incongruent trials the target arrow is flanked by arrows that point in the opposite direction to the target. A priming signal consisting of a fixation cross (500 ms duration) precedes the imperative stimulus (1000 ms duration). After the stimulus, the screen turns blank for 1500 ms. The total inter-stimulus interval is thus 3000 ms. Children were requested to press the right or left key corresponding to the direction of the target arrow as fast and accurately as possible. The task commenced with one practice block of 15 trials (5 trials per condition). This was followed by one experimental block consisting of 60 trials (20 trials for each condition) with the three types of trials presented in a fixed pseudorandom sequence.

Data Analysis

For all tasks, responses below 100 ms were considered anticipatory errors and were excluded from further analyses. Tasks with more than 15% omissions were excluded as described previously (Di Martino et al., 2008). For the two tasks with inhibition, only the go-trials were used (Johnson et al., 2007). Raw time-series data from the TEC and EFT were obtained and plotted to detect possible outliers, data-entry errors or impossible responses (Castellanos et al., 2005). The remaining trials were used to compute mean and coefficient of variation (SD/mean), for each subject and each task, together with accuracy (number of correct responses), number of omissions, and number of directional errors (incorrect responses on the go-trials). First, as initial exploratory analyses, the relationships between RTCV and ratings were assessed via Pearson correlations, which yield the crude, unadjusted associations between measures. Second, to test the primary hypothesis that ISV is associated with parent ratings related to ADHD, RTCV during each task was used as a predictor for the selected subscale after adjusting for age and sex. This step was repeated for the exploratory examination of all the BRIEF subscales. Based on these analyses we quantified the association between RTCV in all tasks and rating subscales in terms of the proportion of the between-subject variance in the subscales explained by between-subject differences in ISV; we report the adjusted $R^2 \times 100$. The same approach was used to assess the contribution of mean RT values and measures of accuracy, omissions, directional errors and commission errors in explaining variability among parent ratings. To test whether RTCV explained a significant proportion of the variance in the ratings after adjusting for measures of accuracy, omissions, or directional errors, we used likelihood ratio tests to compare two models. We compared models with only accuracy, omissions or directional errors (in addition to age and sex) as predictors of rating subscales, to models that also included RTCV as a predictor.

In all analyses, to account for potential correlations between the outcomes of children from the same family, we used models for correlated data. Mixed effects models analyses were performed using the function *lmer* in R (R Development Core Team, 2003). To account for multiple comparisons, we applied Bonferroni corrections. For each task, the significance of the relationship with the five CPRS subscales was judged at $\alpha = 0.05/5 = 0.01$; the significance of the relationship with the 11 BRIEF subscales was judged at $\alpha = 0.05/11 = 0.0045$. We planned to enroll 100 participants to have sufficient statistical power to detect at least medium effect sizes, while adequately controlling for multiple comparisons.

After examining the relationships between measures of performance during each task and the rating subscales, we were interested in determining whether the tasks differed significantly in any particular characteristic. Comparisons between tasks with respect to mean RT, RTCV, accuracy, omissions and directional errors were based on MANOVA-type mixed effects models adjusting for age and sex. Significant task effects (with 4 degrees of freedom, since 5 tasks were compared) were followed by pairwise comparisons between tasks with Bonferroni correction. To test whether RTCV during each task explained

significant proportions of measures of accuracy, omissions and directional errors, we regressed each of these measures on RTCV, adjusting for age and sex. Since the accuracy measures are counts and thus likely to have non-normal distributions, we modeled them using generalized linear models with Poisson regressions.

RESULTS

Group Characteristics

As expected, most children were Latinos ($n = 88$; 90%), the remaining were Caucasian ($n = 6$; 6%), Black ($n = 2$; 2%) or Asian ($n = 2$; 2%). Sixty-eight parents (69%) completed the questionnaires in Spanish; the rest in English. Parents reported that 51 children (52%) had no history of emotional or behavioral difficulties and had not been diagnosed with a psychiatric disorder. Although diagnostic assessments were not conducted, parents of 25 children (26%) reported a past or current diagnosis of ADHD. A total of 14.3% children were reported to be undergoing treatment with psychotropic medications. Eleven children (11%) were taking stimulants (7 methylphenidate and 4 amphetamines), which were withheld on the day of testing; four children (4%) were being treated with neuroleptics. Mean T-scores for the CPRS global DSM-IV subscales varied from 55 to 61 (SD from 12 to 14). Mean T-scores on the BRIEF composite subscales were 55 (SD=14) for the *Behavior Regulation* subscale, 54 (SD=13) for the *Metacognition* subscale and 54 (SD=14) for the *Global Executive Composite* subscale. The BRIEF data of one child were excluded due to excessive inconsistency on the corresponding scale (BRIEF Inconsistency score = 9).

Task Performance

Omission rates exceeded 15% (Di Martino et al., 2008) in one 0B, two 0BI, two 1B, five 1BI and 14 EFT tasks; the data for participants on those tasks were excluded. Additionally, EFT data for another subject were excluded because he only provided five correct responses and 55 errors.

Table 1 displays subjects' performance on the five computer tasks, characterized by the means of reaction time (RTM), reaction time coefficients of variation (RTCV), and the proportions of accurate responses, omissions and directional errors. The mean number of commission errors on the 0BI task was 5 (SD=3) and on the 1BI task was 4 (SD=4).

We formally compared the five computer tasks with respect to subjects' performance. There were only a few differences among tasks, noticeably the small RTCV associated with the EFT, which was significantly lower than on the four TEC tasks. We also studied the associations between performance measures across tasks. After adjusting by age and sex, RTCV was significantly related to measures of accuracy, omissions and directional errors (data not shown).

Associations Between Rating Scales and RTCV

Table 2 shows the Pearson correlations between RTCV measures and parent ratings on the CPRS and BRIEF scales; these correlations are measures of the unadjusted associations between variables. The RTCV measure most consistently associated with the CPRS subscales was 0BI RTCV, which showed significant correlations with all the selected CPRS subscales, ranging from $r=0.33$ to $r=0.39$ ($p<0.01$). Significant correlations were also found between 0B RTCV and *Cognitive Problems/Inattention* ($r=0.26$) and between 1B RTCV and *Hyperactivity* ($r=0.29$) and *DSM-IV: Hyperactive-Impulsive* ($r=0.31$) subscales. None of the CPRS subscales correlated significantly with 1BI or EFT RTCV. Regarding the BRIEF subscales, 0BI RTCV correlated significantly with four subscales (*Inhibit*, *Emotional Control*, *Initiate*, and *Monitor*), two composite scales (*Behavioral Regulation*, and

Metacognition) and with the summary *Global Executive Composite* scale ($r=0.28$ to $r=0.34$; $p<0.0045$). None of the BRIEF subscales correlated significantly with RTCV on the EFT, 0B, 1B or 1BI tasks.

RTCV as Predictor of Rating Scales

The proportion of variance in the rating scales explained by RTCV, after adjusting for age and sex, is reported in Table 3 for each task. The pattern observed in Table 2 is generally preserved, indicating that age and sex do not account for the associations observed in Table 3. RTCV during the 0BI task explained significant proportions of the variance in all the CPRS selected subscales, ranging from 7% to 11% ($p<0.01$). The highest proportion explained was for the *Cognitive Problems/Inattention* subscale (11%). RTCV during the 1B task explained significant proportions of the variance in all but *DSM-IV: Inattentive*, ranging from 4% to 7%. During 0B, 1BI and EFT, RTCV was not related to any of the CPRS subscales related to ADHD.

A similar pattern was observed across the BRIEF subscales. RTCV during the 0BI task explained significant proportions of the variance in four primary BRIEF subscales (*Inhibit*, *Emotional Control*, *Initiate* and *Monitor*), ranging from 4% to 8% ($p<0.0045$), and in all the composite subscales: *Behavioral Regulation* (4%), *Metacognition* (5%) and *Global Executive Composite* (5%). The RTCV during 1B explained a significant proportion of the variance in the *Inhibit* (5%) and *Monitor* (8%) subscales. None of the BRIEF subscales was significantly associated to RTCV for the 0B, 1BI or EFT data sets.

RT Mean, Accuracy, Omissions, Directional and Commission Errors as Predictors

Mean RT was not related to any CPRS subscale nor to any BRIEF subscales for any of the computer tasks (See Table 3). Table 4 shows the percent variance in ratings explained by each of the measures of accuracy, omission, and directional errors, after adjusting for age and sex. For the 0B task, the number of correct responses (accuracy) explained a significant proportion of variance in one CPRS subscale (*Cognitive Problems/Inattention*; 6%) and the number of omissions explained a significant proportion of variance in four CPRS subscales and one BRIEF subscale (ranging from 7% to 11%). For the 0BI task, accuracy explained a significant proportion of variance in two CPRS subscales (5% and 6%) as number of omissions did in one CPRS subscale (5%). For the 1B task, accuracy explained a significant proportion of variance in two CPRS subscales and three BRIEF subscales (ranging from 3% to 5%); omissions explained three CPRS subscales and one BRIEF subscale (ranging from 4% to 7%) and the number of directional errors explained one CPRS subscale and two BRIEF subscales (3% to 4%). Number of omissions on 1BI significantly explained the variance in three CPRS subscales. Measures of accuracy, omissions or directional errors on the EFT did not explain a significant proportion of variance in any rating subscale.

We next asked whether RTCV contributes to explaining variance in the rating scales, even after accounting for the effects of the other performance measures. This was particularly relevant for measures that predicted subscales, such as accuracy during 1B and omissions during 1BI. Underlined entries in Table 4 indicate that RTCV significantly contributed to explaining variance in the rating subscale in that row after adjusting for the performance measure in that column. Consider for example the underlined “4” in the first row and 4th column of Table 4. Accuracy during 1BI explained 4% of the variance in *Cognitive Problems/ Inattention*, which is not statistically significant. From Table 3, RTCV during 1BI is not a significant predictor of *Cognitive Problems/Inattention*. However, after adjusting for accuracy (in addition to age and sex), RTCV during 1BI became a significant predictor. Surprisingly, RTCV during 1BI became a significant predictor for all Conners measures, after adjusting for accuracy, omission or directional errors.

For the 1BI task, adding RTCV to measures of accuracy or directional errors as predictors (adjusting for age and sex), also significantly increased the proportion of variance explained in all the BRIEF subscales. Adding RTCV during 1BI to number of omissions showed the same effect for all five CPRS subscales and four of the BRIEF subscales. For the 0BI task, including RTCV and omissions as predictors increased the proportion of variance explained in all five CPRS and seven BRIEF subscales. In contrast, when the subscales were modeled as a function of RTCV during 0BI along with accuracy or directional errors, the proportion of variance increased in only three of the CPRS subscales. That is, when adjusting for accuracy or omissions, RTCV during 0BI was no longer significant in explaining the variance in two CPRS subscales (*Hyperactivity* and *DSM IV: Hyperactive-Impulsive*) or in the BRIEF subscales. The proportion of variance explained in the subscales did not significantly increase when adding RTCV during 0B, 1B or EFT as predictors of measures of accuracy, omissions or directional errors.

DISCUSSION

In a convenience sample of 98 mostly Latino children, we confirmed our principal hypothesis that parent ratings on ADHD-related Conners scales would be significantly explained by intra-subject variability (ISV) as indexed by response time coefficients of variation (RTCV) on several brief computerized Tasks of Executive Control (TEC). Specifically, 0BI (a simple choice response time task with an inhibitory Go/No Go component) RTCV explained significant proportions of all five Conners subscales and six BRIEF subscales. Similarly, RTCV on the 1B (a one-back task) explained four Conners and three BRIEF subscales. Interestingly, one of the tasks for which we did not observe direct relationships between RTCV and parent ratings (1BI – a one back task requiring inhibition), showed significant increases in the variance explained on all the CPRS subscales when accuracy, omissions or directional errors were included in the models, and on all of the BRIEF subscales when accuracy or directional errors were included. Thus, for three of the four TEC tasks, response time inconsistency was either directly or indirectly related to most parent ratings of ADHD and executive dysfunction.

One of the TEC tasks did not yield consistent relationships. RTCV on the first task presented, 0B (simple choice response time task), showed no significant direct relationships to parent ratings; this was not improved by accounting for performance measures. RTCV during the Eriksen Flanker Task (EFT), which was presented last, also did not exhibit significant relationships with any subscale.

The relatively robust findings for RTCV on several TEC tasks were in contrast to mean RT, which did not explain significant proportions of the variance in any parent subscale or any task, as we had predicted. Our third hypothesis, that RTCV would continue to be significantly related to parent ratings even after adjusting for standard performance measures, was partially supported for the two tasks that also require inhibition (0BI and 1BI).

Overall, our results confirm previous studies that found a positive correlation between increased ISV and ADHD (Klein et al., 2006; Castellanos et al., 2005; Bellgrove et al., 2005; Hervey et al., 2006; Leth-Steensen et al., 2000; Hynd et al., 1989) at least with the TEC tasks that required more than a simple choice (i.e., 0BI, 1B, and 1BI). In a recent comparison between ‘simple’ and ‘complex’ versions of a Go/No-go task, RTCV was significantly greater on the more complex task in the ADHD group compared to controls (Vaurio et al., 2009). In a recent brain imaging study, greater task difficulty was associated with greater ISV and more pronounced differences in functional deactivation in subjects with ADHD compared to controls (Fassbender et al., 2009). Thus, detecting differences

from controls in the variability of response times in ADHD is likely improved when tasks require manipulations beyond a simple choice.

Contrary to our expectations, we did not observe significant relationships between behavior ratings and RTCV on the EFT. This divergence from our past experience (Scheres et al., 2001; Di Martino et al., 2008) might be due to using a shorter version of the task (3 min versus 15-18 min). Since the EFT was always administered at the end of the test session, we cannot determine whether the brevity of this version of the task or the presentation order accounted for our negative EFT results in this sample. Also, the Eriksen Flanker version we used was visually impoverished compared to the TEC tasks which preceded it. We note that we excluded a significantly larger proportion of data on the EFT compared to all the other TEC tasks because of missing data ($X_{(1)}^2 > 5.57$, $p < 0.02$). The mean RT for the remaining subjects was still significantly longer than on the other TEC tasks, while RTCV was significantly lower. We speculate that many of the children who were most variable on the TEC tasks “ran out of gas” by the time they were asked to perform the Flanker task.

Our negative results with the 0B task are consistent with another study that also used an equally brief simple choice response task and failed to find a significant relationship between ISV and ADHD symptoms (Geurts et al., 2008). Future work with versions of these tasks that differ in duration, presented in counterbalanced order, will be needed to define the task parameter ‘boundaries’ required to best observe the expected relationship between ISV and ADHD symptoms.

Limitations of this pilot study include our exclusive dependence on parent report and the lack of teacher’s or clinician’s ratings. The study was intended to test the feasibility of an approach that can be incorporated into large-scale molecular genetic studies. We also wanted to include minority populations that are typically underrepresented. For both these reasons, we limited the time requirement to a single short visit and conducted the study in English or Spanish. These modifications allowed us to enrol a substantial sample of mostly Latino participants who have not been previously examined with regard to ISV and ADHD. We included children who were being treated with stimulants ($n=11$); they were unmedicated for one day prior to testing, as is commonly done. Four children taking atypical neuroleptics were included and not asked to discontinue those medications for ethical reasons. Confirmatory analyses (data not shown) demonstrated the same overall pattern of results even when those four children were excluded. Our sample included about twice as many boys as girls. We adjusted statistically for the effects of sex, which were negligible. Thus, we do not believe that the male preponderance affected our results. Our most notable limitation, which was integral to our low-cost dimensional strategy, was the lack of formal diagnoses derived from structured psychiatric interviews. Despite this intentional absence, we found significant predicted dimensional relationships between ADHD-related subscale scores and ISV, with up to 11% of the variance accounted for, even after adjusting for the considerable effects of age, using brief easily administered computer tasks and rating scales.

In conclusion, children with higher scores on ADHD-related subscales of the CPRS and on the BRIEF subscales exhibited greater RT variability on several relatively brief Executive Control tasks. Overall, variability indices were better at capturing significant relationships to parent ratings than the usual measures of mean response time, accuracy or specific errors. However, despite statistical significance, the relationships with variability we detected are still insufficient in terms of magnitude to be clinically useful. Future work including counterbalanced presentations of fewer than five tasks will be needed to determine their optimal combination. Nevertheless, these initial results support the utility of even these brief assessments of response time variability. Objective variability indices represent a neuropsychological correlate of ADHD that has been hypothesized to reflect underlying

intrinsic fluctuations in large-scale neuronal circuits such as the brain's default-mode network (Castellanos et al., 2005; Sonuga-Barke & Castellanos, 2007). As such, these results bring us closer to developing a mechanistic understanding of the pathophysiology underlying the symptoms of ADHD.

Acknowledgments

The authors appreciate the support of the Alicia Koplowitz Foundation, the Stavros Niarchos Foundation, the Leon Levy Foundation and NIMH (R01MH081218) and the staff members of the Roberto Clemente Family Therapy Center and the pediatrics clinics of Bellevue Hospital Center, New York.

Biographies

Lorena Gómez-Guerrero, M.D., is a child and adolescent psychiatrist trained through postdoctoral fellowships at hospitals in England sponsored by the Castilla del Pino Foundation, and in Spain and the United States sponsored by the Alicia Koplowitz Foundation. She currently works as an attending psychiatrist in an outpatient clinic in Orense, Spain. Her research interests include ADHD, autism, and pharmacogenetics; she also specializes in child and adolescent forensic evaluations.

Cristina Domínguez Martín, M.D., Ph.D., is a child and adolescent psychiatrist at the Hospital Clínico Universitario, Valladolid, Spain where she is an attending psychiatrist in the Day Hospital. She was a clinical and research fellow in child and adolescent psychiatry at the New York University Child Study Center, sponsored by the Alicia Koplowitz Foundation. Her research interests include ADHD, autism and diagnostic assessments.

María Angeles Mairena, Cl.Psy., is a clinical psychologist and research fellow in child and adolescent clinical psychology at the New York University Child Study Center, sponsored by the Alicia Koplowitz Foundation. She is a Ph.D. student at the Universidad Autónoma in Madrid, Spain. Her research interests include ADHD and autism.

Adriana Di Martino, M.D., is research assistant professor of child and adolescent psychiatry at the Phyllis Green and Randolph C. Owen Institute for Pediatric Neuroscience at the New York University Child Study Center. Dr. Di Martino's research focuses on the neurobiology of autism and ADHD.

Jing Wang, M.S., is an assistant research scientist in the department of child and adolescent psychiatry at the New York University Langone Medical Center. Her research interests include family-based interventions for asthma, and autism in toddlers.

Alan L. Mendelsohn, M.D., is associate professor of pediatrics at New York University School of Medicine and Bellevue Hospital Center, New York, NY, where he serves as director of clinical research for the divisions of general and developmental-behavioral pediatrics. His research interests include parent-child interactions, child development, school readiness and electronic media exposure.

Benard P. Dreyer, M.D., is professor of pediatrics at New York University School of Medicine. He is director of developmental-behavioral pediatrics at NYU, and co-investigator on grants examining early childhood development and health literacy. He is the author of a recent review of ADHD in preschool children. His research interests include ADHD, early childhood development and interventions to improve outcomes in high risk populations, effects of media on development and behavior, and health literacy.

Peter K. Isquith, Ph.D., is a pediatric neuropsychologist in independent practice and adjunct assistant professor in psychiatry and pediatrics at Dartmouth Medical School. His research interests include the development and disorders of executive function.

Gerard Gioia, Ph.D., is a pediatric neuropsychologist, chief of the division of pediatric neuropsychology, and director of the Safe Concussion Outcome, Recovery & Education (SCORE) Program at Children's National Medical Center. He is an associate professor of pediatrics and psychiatry at the George Washington University School of Medicine. Dr. Gioia directs an active sports concussions program, serving athletes ranging in age and skill from youth, high school, and college to professional levels. Currently, he is conducting research funded by the Centers for Disease Control to develop effective clinical tests for mild traumatic brain injury/concussion in children as young as 4 years of age.

Eva Petkova, Ph.D., is director of the division of biostatistics at the department of child and adolescent psychiatry at New York University Langone Medical Center and a senior research scientist at the Nathan Klein Institute for Psychiatric Research in Orangeburg, NY.

F. Xavier Castellanos, M.D., Brooke and Daniel Neidich professor of child and adolescent psychiatry and radiology at New York University Langone Medical Center, directs the Phyllis Green and Randolph C. Owen Institute for Pediatric Neuroscience at the NYU Child Study Center and is a senior research scientist at the Nathan Klein Institute for Psychiatric Research in Orangeburg, NY. Dr. Castellanos' research focuses on the neurobiology of ADHD and on identifying quantitative biomarkers for ongoing neuroimaging and molecular genetic studies in child and adolescent psychiatry. Dr. Castellanos co-chairs the DSM-V workgroup on ADHD and Disruptive Behavior Disorders.

References

- Anderson VA, Anderson P, Northam E, Jacobs R, Mikiewicz O. Relationships between cognitive and behavioral measures of executive function in children with brain disease. *Child Neuropsychol.* 2002; 8:231–240. [PubMed: 12759820]
- Andreou P, Neale BM, Chen W, Christiansen H, Gabriels I, Heise A, et al. Reaction time performance in ADHD: improvement under fast-incentive condition and familial effects. *Psychol. Med.* 2007; 37:1–13.
- Bauermeister JJ, Canino G, Bravo M, Ramirez R, Jensen PS, Chavez L, et al. Stimulant and psychosocial treatment of ADHD in Latino/Hispanic children. *Journal of the American Academy of Child and Adolescent Psychiatry.* 2003; 42:851–855. [PubMed: 12819445]
- Bellgrove MA, Hawi Z, Kirley A, Gill M, Robertson IH. Dissecting the attention deficit hyperactivity disorder (ADHD) phenotype: sustained attention, response variability and spatial attentional asymmetries in relation to dopamine transporter (DAT1) genotype. *Neuropsychologia.* 2005; 43:1847–1857. [PubMed: 16168728]
- Castellanos FX. Toward a pathophysiology of attention-deficit/hyperactivity disorder. *Clinical Pediatrics.* 1997; 36:381–393. [PubMed: 9241475]
- Castellanos FX, Sonuga-Barke EJS, Milham MP, Tannock R. Characterizing cognition in ADHD: beyond executive dysfunction. *Trends Cogn Sci.* 2006; 10:117–123. [PubMed: 16460990]
- Castellanos FX, Sonuga-Barke EJS, Scheres A, Di Martino A, Hyde C, Walters JR. Varieties of attention-deficit/hyperactivity disorder-related intra-individual variability. *Biological Psychiatry.* 2005; 57:1416–1423. [PubMed: 15950016]
- Castellanos FX, Tannock R. Neuroscience of attention-deficit hyperactivity disorder: the search for endophenotypes. *Nature Reviews Neuroscience.* 2002; 3:617–628.
- Conners, CK. *Conners' Rating Scales-Revised Technical Manual.* Multi-Health Systems, Inc.; North Tonawanda, NY: 1997.

- Di Martino A, Ghaffari M, Curchack J, Reiss P, Hyde C, Vannucci M, et al. Decomposing intra-subject variability in children with attention-deficit/hyperactivity disorder. *Biological Psychiatry*. 2008; 64:607–614. [PubMed: 18423424]
- Fassbender C, Zhang H, Buzy WM, Cortes CR, Mizuiri D, Beckett L, et al. A lack of default network suppression is linked to increased distractibility in ADHD. *Brain Research*. 2009; 1273:114–128. [PubMed: 19281801]
- Geurts HM, Grasman RPPP, Verte S, Oosterlaan J, Roeyers H, van Karmmen SM, et al. Intra-individual variability in ADHD, autism spectrum disorders and Tourette's syndrome. *Neuropsychologia*. 2008; 46:3030–3041. [PubMed: 18619477]
- Gioia, GA.; Isquith, PK.; Guy, SC.; Kenworthy, L. Behavior Rating Inventory of Executive Function. Psychological Assessment Resources, Inc.; Lutz, FL: 2000.
- Gioia GA, Isquith PK, Kenworthy L, Barton RM. Profiles of everyday executive function in acquired and developmental disorders. *Neuropsychol.Dev Cogn C Child Neuropsychol*. 2002; 8:121–137.
- Happé F, Booth R, Charlton R, Hughes C. Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: Examining profiles across domains and ages. *Brain and Cognition*. 2006; 61:25–39. [PubMed: 16682102]
- Hervey AS, Epstein JN, Curry JF, Tonev S, Eugene AL, Keith CC, et al. Reaction time distribution analysis of neuropsychological performance in an ADHD sample. *Child Neuropsychol*. 2006; 12:125–140. [PubMed: 16754533]
- Hynd GW, Nieves N, Connor RT, Stone P, Town P, Becker MG, et al. Attention deficit disorder with and without hyperactivity: reaction time and speed of cognitive processing. *Journal of Learning Disabilities*. 1989; 22:573–580. [PubMed: 2809410]
- Isquith, PK.; Roth, RM.; Gioia, GA. Tasks of Executive Control (TEC). Psychological Assessment Resources, Inc.; Lutz, FL: 2009.
- Johnson KA, Kelly SP, Bellgrove MA, Barry E, Cox M, Gill M, et al. Response variability in attention deficit hyperactivity disorder: evidence for neuropsychological heterogeneity. *Neuropsychologia*. 2007; 45:630–638. ePub 12/8/2006. [PubMed: 17157885]
- Klein C, Wendling K, Huettner P, Ruder H, Peper M. Intra-subject variability in attention-deficit hyperactivity disorder (ADHD). *Biological Psychiatry*. 2006; 60:1088–1097. [PubMed: 16806097]
- Kuntsi J, Rogers H, Swinard G, Borger N, Meere JV, Rijdsdijk F, et al. Reaction time, inhibition, working memory and 'delay aversion' performance: genetic influences and their interpretation. *Psychol.Med*. 2006:1–12.
- Kuntsi J, Stevenson J. Psychological mechanisms in hyperactivity: II The role of genetic factors. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 2001; 42:211–219.
- Lahey BB, Applegate B, McBurnett K, Biederman J, Greenhill L, Hynd GW, et al. DSM-IV field trials for attention deficit hyperactivity disorder in children and adolescents. *American Journal of Psychiatry*. 1994; 151:1673–1685. [PubMed: 7943460]
- Leth-Steensen C, Elbaz ZK, Douglas VI. Mean response times, variability, and skew in the responding of ADHD children: a response time distributional approach. *Acta Psychologica (Amsterdam)*. 2000; 104:167–190.
- Levy F, Hay DA, McStephen M, Wood C, Waldman I. Attention-deficit hyperactivity disorder: a category or a continuum? Genetic analysis of a large-scale twin study. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1997; 36:737–744. [PubMed: 9183127]
- Nigg JT. Neuropsychologic theory and findings in attention-deficit/hyperactivity disorder: the state of the field and salient challenges for the coming decade. *Biological Psychiatry*. 2005; 57:1424–1435. [PubMed: 15950017]
- R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2003.
- Rosvold HE, Mirsky AF, Sarason I, Bransome ED, Beck LH. A continuous performance test of brain damage. *Journal of Consulting Psychology*. 1956; 20:343–350. [PubMed: 13367264]
- Scheres A, Oosterlaan J, Sergeant JA. Response execution and inhibition in children with AD/HD and other disruptive disorders: the role of behavioural activation. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 2001; 42:347–357.

- Sonuga-Barke EJ, Castellanos FX. Spontaneous attentional fluctuations in impaired states and pathological conditions: A neurobiological hypothesis. *Neuroscience and Biobehavioral Reviews*. 2007; 31:977–986. [PubMed: 17445893]
- Sonuga-Barke EJS, Sergeant JA, Nigg J, Willcutt E. Executive dysfunction and delay aversion in attention deficit hyperactivity disorder: Nosologic and diagnostic implications. *Child and Adolescent Psychiatric Clinics of North America*. 2008; 17:367–384. [PubMed: 18295151]
- Toplak ME, Bucciarelli S, Jain U, Tannock R. Executive Functions: Performance-Based Measures and the Behavior Rating Inventory of Executive Function (BRIEF) in Adolescents with Attention Deficit/Hyperactivity Disorder (ADHD). *Child Neuropsychology*. 2008; 15:53–72. [PubMed: 18608232]
- Vaurio RG, Simmonds DJ, Mostofsky SH. Increased intra-individual reaction time variability in attention-deficit/hyperactivity disorder across response inhibition tasks with different cognitive demands. *Neuropsychologia*. 2009
- Willcutt EG, Doyle AE, Nigg JT, Faraone SV, Pennington BF. Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biological Psychiatry*. 2005; 57:1336–1346. [PubMed: 15950006]

Table 1

Task Performance

	MRT (ms) Mean (SD)	RTCV Mean (SD)	% Accuracy Mean (SD)	% Omissions Mean (SD)	% Directional Errors Mean (SD)
0B	457.1 (104.6)	0.29 (0.08)	87% (12%)	1% (2%)	11% (10%)
0BI	496.8 (113.4)	0.29 (0.10)	88% (11%)	2% (4%)	10% (8%)
1B	509.6 (151.3)	0.30 (0.09)	84% (11%)	2% (2%)	14% (9%)
1BI	528.6 (153.2)	0.32 (0.10)	82% (12%)	2% (2%)	14% (8%)
EFT	561 (84.7)	0.22 (0.05)	87% (8%)	3% (5%)	7% (7%)

0B: Zero-Back, a simple choice response task with a designated target; 0BI: Zero-Back with an inhibitory component. 1B: 1-Back; 1BI: 1-Back with an inhibitory component; EFT: Eriksen Flanker Task; MRT: Mean Response Time on Go Trials; SD: Standard Deviation; RTCV: Response Time Coefficient of Variation

Table 2Pearson correlations (r) between parent rating scales and response time coefficient of variation (RTCV)

	0B- RTCV	0BI - RTCV	1B - RTCV	1BI - RTCV	EFT - RTCV
<i>CPRS</i>					
B. Cognitive Problems/Inattention	0.26 *	0.39 *	0.19	0.19	0.12
C. Hyperactivity	0.17	0.35 *	0.29 *	0.21	0.18
L. DSM-IV: Inattentive	0.17	0.36 *	0.21	0.22	0.08
M. DSM-IV: Hyperactive- Impulsive	0.19	0.33 *	0.31 *	0.22	0.24
N. DSM-IV: Total	0.18	0.34 *	0.26	0.22	0.17
<i>BRIEF</i>					
Inhibit	0.12	0.33 *	0.29	0.21	0.19
Shift	0.11	0.29	0.15	0.16	0.20
Emotional Control	0.02	0.33 *	0.17	0.18	0.09
Behavioral Regulation Composite	0.10	0.34 *	0.23	0.20	0.16
Initiate	0.13	0.33 *	0.22	0.18	0.03
Working Memory	0.12	0.22	0.22	0.13	0.05
Plan/Organize	0.09	0.21	0.21	0.16	0.06
Organization of Materials	0.06	0.17	0.04	0.13	0.03
Monitor	0.13	0.32 *	0.27	0.16	0.13
Metacognition Composite	0.08	0.28 *	0.16	0.14	0.01
Global Executive Composite	0.11	0.31 *	0.21	0.17	0.10

CPRS: Conners' Parent Rating Scale-Revised: Long Version; BRIEF: Behavior Rating Inventory of Executive Function; RTCV: Response Time Coefficient of Variation. 0B: Zero-Back or simple choice response task with a designated target; 0BI: Zero-Back with an inhibitory component; 1B: 1-Back; 1BI: 1-Back with an inhibitory component; EFT: Eriksen Flanker Task.

* Significant p-values at $p < 0.01$ for CPRS subscales and at $p < 0.0045$ for BRIEF subscales.

Table 3

Percentage variance explained ($100 \times R^2$) in each rating subscale by mean of response time (RTM) and coefficient of variation of response time (RTCV) in OB, OBI, IB, IBI, and EFT, after adjusting for age and sex and accounting for correlations between family members.

	RTM					RTCV				
	OB	OBI	IB	IBI	EFT	OB	OBI	IB	IBI	EFT
CPRS										
B. Cognitive Problems/Inattention	1	4	3	3	2	3	11*	7*	2	1
C. Hyperactivity	0	0	0	0	1	1	7*	5*	3	2
L. DSM-IV: Inattentive	0	2	1	1	1	1	9*	2	3	1
M. DSM-IV: Hyperactive-Impulsive	0	0	0	0	2	2	7*	7*	4	4
N. DSM-IV: Total	0	1	0	1	1	1	7*	4*	3	2
BRIEF										
Inhibit	0	0	0	0	0	1	6*	5*	3	3
Shift	0	0	0	0	0	0	2	1	1	1
Emotional Control	0	0	0	0	0	0	5*	2	2	0
Behavioral Regulation	0	0	0	0	0	0	4*	3	2	1
Initiate	1	1	2	2	1	1	5*	3	2	0
Working Memory	2	4	2	3	1	0	3	3	1	0
Plan/Organize	0	2	2	3	2	0	6	4	3	0
Organization of Materials	1	2	3	3	1	0	2	0	1	0
Monitor	0	2	1	1	1	1	8*	6*	2	2
Metacognition	0	2	2	2	1	0	5	1	2	0
Global Executive Composite	0	1	1	1	1	0	5*	3	2	1

CPRS: Conners' Parent Rating Scale-Revised: Long Version; BRIEF: Behavior Rating Inventory of Executive Function; RT: Response Time; OB: Zero-Back or simple choice response; OBI: Zero-Back with an inhibitory component; IB: 1-Back; IBI: 1-Back with an inhibitory component; EFT: Eriksen Flanker Task.

* Significant p-values at $p < 0.01$ for CPRS subscales and at $p < 0.0045$ for BRIEF subscales.

Table 4

Percentage variance explained ($100 \times R^2$) in each rating subscale by measures of accuracy, omissions and directional errors in OB, OBI, 1B, 1BI, and EFT, after adjusting for age, sex and accounting for correlations between family members; and comparison with regression models adjusting for response time coefficient of variation.

	Accuracy Explaining Subscales					Omissions Explaining Subscales					Directional Errors Explaining Subscales				
	OB	OBI	1B	1BI	EFT	OB	OBI	1B	1BI	EFT	OB	OBI	1B	1BI	EFT
CPRS															
B. Cognitive Problems/Inattention	6*	4	5*	4	0	11*	1	7*	7*	0	3	3	3	3	1
C. Hyperactivity	2	5*	4*	0	1	4	1	3	2	0	1	5*	4*	2	1
L. DSM-IV: Inattentive	3	3	3	2	0	8*	1	4	5*	0	1	2	1	1	0
M. DSM-IV: Hyperactive-Impulsive	3	6*	4	1	1	7*	1	5*	4	0	2	5*	2	1	0
N. DSM-IV: Total	3	3	3	1	0	7*	1	4*	4*	0	1	3	2	1	0
BRIEF															
Inhibit	3	5	5*	1	0	5*	1	5*	3	0	2	5	4*	2	0
Shift	0	0	1	1	0	1	0	0	2	1	0	0	1	1	0
Emotional Control	1	1	3*	0	0	1	0	2	1	0	0	1	2	1	0
Behavioral Regulation	1	2	3*	0	0	2	1	2	3	0	1	2	3*	1	0
Initiate	1	1	3	3	0	2	0	2	2	0	1	1	2	1	0
Working Memory	1	1	1	0	0	3	0	1	2	0	0	1	0	0	0
Plan/Organize	1	1	2	2	0	6	0	2	4	0	0	0	1	1	0
Organization of Materials	0	0	0	0	1	0	1	0	2	0	0	0	0	0	0
Monitor	3	2	3	1	0	5	0	3	3	0	2	2	1	2	0
Metacognition	1	0	1	1	0	3	0	1	2	0	0	0	0	0	0
Global Executive Composite	1	1	2	1	0	3	0	2	3	0	1	1	2	1	0

CPRS: Conners' Parent Rating Scale-Revised; Long Version; BRIEF: Behavior Rating Inventory of Executive Function; OB: Zero-Back or simple choice response task with a designated target; OBI: Zero-Back with an inhibitory component; 1B: 1-Back; 1BI: 1-Back with an inhibitory component; EFT: Eriksen Flanker Task.

* Significant p-values at $p < 0.01$ for CPRS subscales and at $p < 0.0045$ for BRIEF subscales; Underlined: Adding the Response Time Coefficient of Variation as a predictor increases significantly the proportion of variance explained in the rating subscale (Significant p-values at $p < 0.01$ for CPRS subscales and at $p < 0.0045$ for BRIEF subscales).